

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 9.51698 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-55
Perfect score: 124
Sequence: 1 EOASVKYVILDMYRALTLMTSTAT 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44.5	35.9	798	4	US-09-861-451A-12
2	44	35.5	278	1	US-07-921-807B-10
3	44	35.5	278	1	US-08-441-944A-10
4	43	34.7	428	4	US-09-328-352-5543
5	42	33.9	301	3	US-09-134-001C-4305
6	42	33.9	332	4	US-09-710-279-2350
7	41	33.1	248	4	US-09-248-796A-15839
8	41	33.1	425	1	US-08-700-749A-5
9	41	33.1	425	3	US-09-020-684-5
10	41	33.1	425	3	US-09-020-467-5
11	41	33.1	425	3	US-09-020-685-5
12	41	33.1	425	3	US-09-020-683-5
13	41	33.1	619	4	US-09-270-767-34769
14	41	33.1	619	4	US-09-270-767-49986
15	40.5	32.7	740	4	US-09-489-039A-11157
16	40	32.3	65	4	US-09-328-352-7222
17	40	32.3	161	2	US-08-403-852D-30
18	40	32.3	161	3	US-08-510-646B-31
19	40	32.3	161	3	US-09-231-818-30
20	40	32.3	161	4	US-09-635-359B-30
21	40	32.3	177	4	US-09-614-912-136
22	40	32.3	305	4	US-09-328-352-6794
23	40	32.3	527	4	US-09-538-092-238
24	40	32.3	531	4	US-09-252-991A-27660
25	39.5	31.9	65	4	US-09-248-796A-27954
26	39.5	31.9	345	4	US-09-270-767-60654
27	39.5	31.9	404	4	US-09-270-767-45164

28	39.5	31.9	998	4	US-09-540-236-3801	Sequence 3801, Ap
29	39.5	31.9	1724	2	US-08-477-451-15	Sequence 15, Appl
30	39	31.5	112	4	US-09-248-796A-26280	Sequence 26280, A
31	39	31.5	178	4	US-09-543-681A-7128	Sequence 7128, Ap
32	39	31.5	236	4	US-09-107-532A-6356	Sequence 6356, Ap
33	39	31.5	305	4	US-09-340-236-3362	Sequence 3362, Ap
34	39	31.5	317	4	US-09-248-796A-20568	Sequence 20568, A
35	39	31.5	370	2	US-08-837-593-7	Sequence 7, Appl1
36	39	31.5	375	2	US-08-837-593-6	Sequence 6, Appl1
37	39	31.5	375	2	US-08-837-593-5	Sequence 5, Appl1
38	39	31.5	375	2	US-09-623-034-2	Sequence 2, Appl1
39	39	31.5	387	2	US-08-837-593-4	Sequence 4, Appl1
40	39	31.5	393	2	US-08-837-593-3	Sequence 3, Appl1
41	39	31.5	431	4	US-09-270-767-42950	Sequence 42950, A
42	39	31.5	434	4	US-09-538-092-209	Sequence 209, App
43	39	31.5	450	4	US-09-721-870-16	Sequence 16, Appl
44	39	31.5	551	4	US-09-710-279-1668	Sequence 1668, Ap
45	39	31.5	560	3	US-09-134-001C-3153	Sequence 3153, Ap

ALIGNMENTS

RESULT 1
US-09-861-451A-12
; Sequence 12, Application US/09861451A
; Patent No. 6759516
; GENERAL INFORMATION:
; APPLICANT: Commonwealth Scientific & Industrial Research Orga
; TITLE OF INVENTION: Methods of Identifying Antigen Gene Sequences
; FILE REFERENCE: PF34033/01
; CURRENT APPLICATION NUMBER: US/09/861,451A
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PF7273
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 798
; TYPE: EXT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Deduced protein
; OTHER INFORMATION: sequence from clone PAD653
US-09-861-451A-12

Query Match 35.9%; Score 44.5; DB 4; Length 798;
Best Local Similarity 39.3%; Pred.No. 76;
Matches 11; Conservative 7; Mismatches 7; Indels 3; Gaps 1;

QY 1 EOASVKYVIL--DMYRALTLMTSTA 25
DB 605 EATKFNILVAESAYQIVSLFNNSNA 632

RESULT 2
US-07-921-807B-10
; Sequence 10, Application US/07921807B
; Patent No. 5474314
; GENERAL INFORMATION:
; APPLICANT: SPAETE, RICHARD
; TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
; TITLE OF INVENTION: OF VIRAL PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 Horton Street - R440
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

; MOLECULE TYPE: protein
US-08-441-944A-10

Query Match      35.5%; Score 44; DB 1; Length 278;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 11 DMYRALTLTLMNSTA 25
Db 117 DQLRALTLTLLSSDTA 131

RESULT 4
US-09-328-352-5543
; Sequence 5543, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5543
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5543

Query Match      34.7%; Score 43; DB 4; Length 428;
Best Local Similarity 42.9%; Pred. No. 65;
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 5 VKYVILDMYRALTLTLMNSTA 25
Db 89 VPYITLDFIHCLLSANNIALA 109

RESULT 5
US-09-134-001C-4305
; Sequence 4305, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4305
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4305

Query Match      33.9%; Score 42; DB 3; Length 301;
Best Local Similarity 30.4%; Pred. No. 63;
Matches 7; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

Qy 2 QASQKVVILDMYRALTLTLMNST 24
Db 273 BEAKHLIFDALTWEVINTGT 295

RESULT 6
US-09-710-279-2350
; Sequence 2350, Application US/09710279

```



```
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P034800S
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 2350
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-2350

Query Match      33.9%; Score 42; DB 4; Length 332;
Best Local Similarity 30.4%; Pred. No. 70;
Matches 7; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY      2 QASVKYVILDMYRALLTMTST 24
      : : : : : : : : : : : : : : : : : : : : : : : :
Db      304 EEAKKHLFDALDAMEVLNTGT 326

RESULT 7
US-09-248-796A-15839
; Sequence 15839, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15839
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15839

Query Match      33.1%; Score 41; DB 4; Length 248;
Best Local Similarity 40.0%; Pred. No. 73;
Matches 10; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY      2 QASVKYVILDMYRALLTMTST 26
      : : : : : : : : : : : : : : : : : : : : : : : :
Db      91 QASLSISAQVLAATATNTATAT 115

RESULT 8
US-08-700-749A-5
; Sequence 5, Application US/08700749A
; Patent No. 5789550
; GENERAL INFORMATION:
; APPLICANT: GOEDEL, DAVID V.
; TITLE OF INVENTION: TRAF INHIBITORS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA way
; CITY: South San Francisco

; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P034800S
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 2350
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-2350

Query Match      33.9%; Score 42; DB 4; Length 332;
Best Local Similarity 30.4%; Pred. No. 70;
Matches 7; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY      2 QASVKYVILDMYRALLTMTST 24
      : : : : : : : : : : : : : : : : : : : : : : : :
Db      304 EEAKKHLFDALDAMEVLNTGT 326

RESULT 7
US-09-248-796A-15839
; Sequence 15839, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15839
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15839

Query Match      33.1%; Score 41; DB 4; Length 248;
Best Local Similarity 40.0%; Pred. No. 73;
Matches 10; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY      2 QASVKYVILDMYRALLTMTST 26
      : : : : : : : : : : : : : : : : : : : : : : : :
Db      91 QASLSISAQVLAATATNTATAT 115

RESULT 8
US-08-700-749A-5
; Sequence 5, Application US/08700749A
; Patent No. 5789550
; GENERAL INFORMATION:
; APPLICANT: GOEDEL, DAVID V.
; TITLE OF INVENTION: TRAF INHIBITORS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA way
; CITY: South San Francisco
```

```
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,749A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/002382
; FILING DATE: 17-aug-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P0960R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 5:
; LENGTH: 425 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-700-749A-5

Query Match      33.1%; Score 41; DB 1; Length 425;
Best Local Similarity 34.8%; Pred. No. 1.4e+02;
Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY      1 EQASVKYVILDMYRALLTMTS 23
      : : : : : : : : : : : : : : : : : : : : : : : :
Db      46 EQLSLQQTIDKLKSLQLLVNST 68

RESULT 9
US-09-020-684-5
; Sequence 5, Application US/09020684
; Patent No. 6004553
; GENERAL INFORMATION:
; APPLICANT: GOEDEL, DAVID V.
; TITLE OF INVENTION: TRAF INHIBITORS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,684
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/002382
; FILING DATE: 17-aug-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P0960R1D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
```

```

; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 425 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-020-684-5

```

```

Query Match      33.1%; Score 41; DB 3; Length 425;
Best Local Similarity 34.8%; Pred. No. 1.4e+02;
Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy 1 EQASVKYVILDMYRALLTLNMTS 23
Db 46 EQLSLOQTIDKLKSQLLLVNST 68

```

RESULT 10

```

US-09-020-467-5
Sequence 5, Application US/09020467

```

```

; Patent No. 6060303
; GENERAL INFORMATION:
; APPLICANT: GOEDDEL, DAVID V.
; APPLICANT: ROTHE, MIKE
; TITLE OF INVENTION: TRAF INHIBITORS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,467
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/002382
; FILING DATE: 17-aug-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P0960R1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 425 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-020-467-5

```

```

Query Match      33.1%; Score 41; DB 3; Length 425;
Best Local Similarity 34.8%; Pred. No. 1.4e+02;
Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy 1 EQASVKYVILDMYRALLTLNMTS 23
Db 46 EQLSLOQTIDKLKSQLLLVNST 68

```

RESULT 11

```

US-09-020-685-5
Sequence 5, Application US/09020685
; Patent No. 6063585
; GENERAL INFORMATION:
; APPLICANT: GOEDDEL, DAVID V.

```

```

; APPLICANT: ROTHE, MIKE
; TITLE OF INVENTION: TRAF INHIBITORS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,685
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/002382
; FILING DATE: 17-aug-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P0960R1D4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 425 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-020-685-5

```

```

Query Match      33.1%; Score 41; DB 3; Length 425;
Best Local Similarity 34.8%; Pred. No. 1.4e+02;
Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy 1 EQASVKYVILDMYRALLTLNMTS 23
Db 46 EQLSLOQTIDKLKSQLLLVNST 68

```

RESULT 12

```

US-09-020-683-5
Sequence 5, Application US/09020683
; Patent No. 6294348
; GENERAL INFORMATION:
; APPLICANT: GOEDDEL, DAVID V.
; APPLICANT: ROTHE, MIKE
; TITLE OF INVENTION: TRAF INHIBITORS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,683
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/002382
; FILING DATE: 17-aug-1995

```

ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: P0960R1D3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-3216
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-09-020-683-5

Query Match 33.1%; Score 41; DB 3; Length 425;
Best Local Similarity 34.8%; Pred. No. 1.4e+02;
Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1 EQASVKYVILDMYRALTLMTS 23
DB 46 EQLSQQTIIDKLSQLLVNST 68

RESULT 13
US-09-270-767-34769
Sequence 34769, Application US/09270767
Patent No. 6703491

GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 34769
LENGTH: 619
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-34769

Query Match 33.1%; Score 41; DB 4; Length 619;
Best Local Similarity 34.5%; Pred. No. 2.1e+02;
Matches 10; Conservative 6; Mismatches 9; Indels 4; Gaps 1;

QY 2 QASVKYVILDMYRAL---LTLMTSTAT 26
DB 418 QGAYDYTYIEIYRYLYIFHTCWTST 446

RESULT 14
US-09-270-767-49986
Sequence 49986, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 49986
LENGTH: 619
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-49986

Query Match 33.1%; Score 41; DB 4; Length 619;

Best Local Similarity 34.5%; Pred. No. 2.1e+02;
Matches 10; Conservative 6; Mismatches 9; Indels 4; Gaps 1;

QY 2 QASVKYVILDMYRAL---LTLMTSTAT 26
DB 418 QGAYDYTYIEIYRYLYIFHTCWTST 446

RESULT 15
US-09-489-039A-11157
Sequence 11157, Application US/09489039A
Patent No. 6610836

GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 11157
LENGTH: 740
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11157

Query Match 32.7%; Score 40.5; DB 4; Length 740;
Best Local Similarity 32.1%; Pred. No. 3.1e+02;
Matches 9; Conservative 7; Mismatches 3; Indels 9; Gaps 1;

QY 3 ASVKYVILDMYRA-----LTLTLMN 21
DB 156 SSVAYIVIDEFHAFISGSRGVQLLSLN 183

Search completed: November 10, 2004, 14:55:40
Job time : 10.567 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 30.0717 Seconds
(without alignments)

305.399 Million cell updates/sec

Title: US-10-092-750-55

Perfect score: 124

Sequence: 1 EQASVKVILDMYRALTLTMTSTAT 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	124	100.0	26	14	Sequence 55, Appl
2	48.5	39.1	398	15	Sequence 185720,
3	48	38.7	357	16	Sequence 136288,
4	47.5	38.3	122	15	Sequence 254656,
5	47	37.9	95	11	Sequence 8064, Ap
6	45	36.3	78	17	Sequence 311380,
7	45	36.3	1205	15	Sequence 2967, Ap
8	44.5	35.9	798	9	Sequence 12, Appl
9	44	35.5	91	15	Sequence 271940,
10	44	35.5	109	15	Sequence 270166,
11	44	35.5	737	15	Sequence 256196,
12	43.5	35.1	307	9	Sequence 6331, Ap
13	43	34.7	103	16	Sequence 38851, A

14	43	34.7	420	15	US-10-282-122A-44699	Sequence 44699, A
15	43	34.7	477	15	US-10-424-599-175572	Sequence 175572,
16	43	34.7	858	15	US-10-425-114-72451	Sequence 72451, A
17	43	34.7	972	15	US-10-424-599-219651	Sequence 219651,
18	42.5	34.3	237	16	US-10-741-131-35	Sequence 35, Appl
19	42.5	34.3	237	16	US-10-742-350-35	Sequence 35, Appl
20	42	33.9	43	15	US-10-424-599-186763	Sequence 186763,
21	42	33.9	43	15	US-10-424-599-212841	Sequence 212841,
22	42	33.9	108	17	US-10-425-115-361202	Sequence 361202,
23	42	33.9	138	17	US-10-425-115-284221	Sequence 284221,
24	42	33.9	206	17	US-10-425-115-284217	Sequence 128472,
25	42	33.9	352	16	US-10-437-963-125472	Sequence 125472,
26	42	33.9	359	17	US-10-425-115-284222	Sequence 284222,
27	42	33.9	380	15	US-10-425-114-48492	Sequence 48492, A
28	42	33.9	380	15	US-10-425-114-68515	Sequence 68515, A
29	42	33.9	380	15	US-10-425-114-70564	Sequence 70564, A
30	42	33.9	1071	16	US-10-437-963-176644	Sequence 176644,
31	41.5	33.5	129	17	US-10-425-115-245548	Sequence 245548,
32	41.5	33.5	231	17	US-10-425-115-328441	Sequence 328441,
33	41	33.1	65	15	US-10-424-599-276237	Sequence 276237,
34	41	33.1	119	15	US-10-424-599-197293	Sequence 197293,
35	41	33.1	119	17	US-10-788-792-189	Sequence 189, App
36	41	33.1	154	17	US-10-425-115-276085	Sequence 276085,
37	41	33.1	154	17	US-10-425-114-51400	Sequence 51400, A
38	41	33.1	227	15	US-10-424-599-284783	Sequence 284783,
39	41	33.1	281	16	US-10-767-701-41939	Sequence 41939, A
40	41	33.1	287	14	US-10-263-568-9	Sequence 9, Appl
41	41	33.1	347	15	US-10-425-114-73116	Sequence 73116, A
42	41	33.1	362	16	US-10-437-963-115286	Sequence 115286,
43	41	33.1	381	14	US-10-369-493-19997	Sequence 19937, A
44	41	33.1	409	14	US-10-369-493-10754	Sequence 10754, A
45	41	33.1	425	14	US-10-247-671-148	Sequence 148, App

ALIGNMENTS

RESULT 1
US-10-092-750-55
; Sequence 55, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-55

Query Match 100.0%; Score 124; DB 14; Length 26;
Best Local Similarity 100.0%; Pred. No. 4.4e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EQASVKVILDMYRALTLTMTSTAT 26
Db 1 EQASVKVILDMYRALTLTMTSTAT 26

RESULT 2
US-10-424-599-185720
; Sequence 185720, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

```

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 39-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 185720
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(398)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_138719C.1.pap
; US-10-424-599-185720

```

```

RESULT 3
US-10-437-963-136288
; Sequence 136288, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 136288
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_37882C.1.pbp
US-10-437-963-136288

```

RESULT 4
US-10-424-599-254656
; Sequence 254656, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J

```

; APPLICANT: Kovalic David K
; APPLICANT: Cao Yongwei
; APPLICANT: Zhou Yihua
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCES: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 254656
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(122)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_71978C.1.pep
; US-10-424-599-254656

```

```

RESULT 5
US-09-864-408A-8064
; Sequence 8064, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1el Human Polynucleotides and Polypeptides Enc
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8064
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
US-09-864-408A-8064

```

RESULT 6
US-10-425-115-311380
; Sequence 311380, Application US/10425115
; Publication No. US20040314272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

	Matches	10;	Conservative	5;	Mismatches	4;	Indels
Qy	1	EQASVKYVLDMYRALITLKN	21				
		:::-- ::					
Db	17	EQSSLGF--LDLYQALLVFFN	35				

	Matches	10;	Conservative	5;	Mismatches	4;	Indels
Qy	1	EQASVKYVLDMYRALITLKN	21				
		:::-- ::					
Db	17	EQSSLGF--LDLYQALLVFFN	35				

RESULT 11

```

US-10-424-599-256196
; Sequence 256196, Application US/10424599
; Publication NO. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Mole
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 256196
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; -OTHER INFORMATION: Clone ID: PAT_MRT3847
US-10-424-599-256196

```

```
Query Match      35.5%; Score 44; DB 15; Length 737;
Best Local Similarity 66.7%; Pred. NO. 2.7e+02;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

QY 10 LDMYRALLTLMNTST 24
||| |||: |||
Db 146 LDMELELTMMNTLT 160

RESULT 12

```

US-09-738-626-6331
; Sequence 6331, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIHO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6331
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6331

```

Query Match 35.1%; Score 43.5; DB 9; Length 307;
Best Local Similarity 44.0%; Pred. No. 1.2e+02;
Matches 11; Conservative 5; Mismatches 4; Indels 5; Gaps 1;

2 QASVKYVILDMYRALLTLMNTSTAT 26
 |||: :|||:| :||
 255 QASLPLMLMDMYKA----GTAPAT 274

RESULT 13

```

US-10-767-701-38851
; Sequence 38851, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 38851
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C76631_1.pep
US-10-767-701-38851

```

Query Match 34.7%; Score 43; DB 16; Length 103;
Best Local Similarity 36.4%; Pred. No. 41;
Matches 8; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy	3 ASVKVILDMYRALLTLMTST 24 : : : : : : :
Db	16 SDLKLVFDIYQPLLDLVQNPT 37

RESULT 14

US-10-282-122A-44699
; Sequence 44699, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Orlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

```

1  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
2
3  FILE REFERENCE: ELITRA.034A
4
5  CURRENT APPLICATION NUMBER: US/10/282,122A
6
7  CURRENT FILING DATE: 2003-02-20
8
9  PRIOR APPLICATION NUMBER: 60/191,078
10
11  PRIOR FILING DATE: 2000-03-21
12
13  PRIOR APPLICATION NUMBER: 60/206,848
14
15  PRIOR FILING DATE: 2000-05-23
16
17  PRIOR APPLICATION NUMBER: 60/207,727
18
19  PRIOR FILING DATE: 2000-05-26
20
21  PRIOR APPLICATION NUMBER: 60/230,335
22
23  PRIOR FILING DATE: 2000-09-06
24
25  PRIOR APPLICATION NUMBER: 60/230,347
26
27  PRIOR FILING DATE: 2000-09-09
28
29  PRIOR APPLICATION NUMBER: 60/242,578
30
31  PRIOR FILING DATE: 2000-10-23
32
33  PRIOR APPLICATION NUMBER: 60/253,625
34
35  PRIOR FILING DATE: 2000-11-27
36
37  PRIOR APPLICATION NUMBER: 60/257,931
38
39  PRIOR FILING DATE: 2000-12-22
40
41  PRIOR APPLICATION NUMBER: 60/267,636
42
43  PRIOR FILING DATE: 2001-02-09
44
45  PRIOR APPLICATION NUMBER: 60/269,308
46
47  PRIOR FILING DATE: 2001-02-16

```



```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44599
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-10-282-122A-44599

Query Match          34.7%; Score 43; DB 15; Length 420;
Best Local Similarity 42.9%; Pred. No. 2.1e+02;
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy      5 VKYVILDMYRALLTILMTSTA 25
Db      81 VPIITLDFHCLLSAMNIALA 101

RESULT 15
US-10-424-599-175572
; Sequence 175572, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 175572
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_12955C.1.pep
US-10-424-599-175572

Query Match          34.7%; Score 43; DB 15; Length 477;
Best Local Similarity 38.5%; Pred. No. 2.4e+02;
Matches 10; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

Qy      1 EQASVKYVILDMYRALLTILMTSTAT 26
Db      182 EQRSAPYSMDIRSAVYTQNNVIGT 207

Search completed: November 11, 2004, 02:43:11
Job time : 31.1217 secs
```

This Page Blank (uspto)

C;Genetics:
A;Gene: sqdB
C;Superfamily: sulfolipid biosynthesis protein sqdB

Query Match 40.3%; Score 50; DB 2; Length 389;
Best Local Similarity 42.3%; Pred. No. 3; 5;
Matches 11; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 EQASVKYVILDMYRALTLMTSTAT 26
|||:::|:::|:::|:::|
Db 99 EORSAPYSMDVNHANYTLINLTST 124
|||:::|:::|:::|:::|

RESULT 3
A86335
T20H2.9 protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Nov-2001
C;Accession: A86335
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Pedersen, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Hughes, B.K.; Conn, L.; Conway, A.S.; Conway, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408: 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; PMID:21016719; PMID:11130712
A;Accession: A86335
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1137 <STO>
A;Cross-references: GB:AEO05172; NID:g8778985; PIDN:AAF79900.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

Query Match 37.9%; Score 47; DB 2; Length 1137;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 6 KYVILDMYRALTLTMN 21
|:::|:::|:::|:::|
Db 1075 KFYYIDMYNSLLEVIN 1090
|:::|:::|:::|:::|

RESULT 4
A99489
arabinose ABC transporter, permease SS03068 [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: A99489
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Avayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to Genbank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: A99489
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-235 <KUR>
A;Cross-references: UNIPROT:Q97UF3; GB:AEO06641; NID:gl3816476; PIDN:AAK43168.1; GSPDB:G
C;Genetics:
A;Gene: SS03068

Query Match 35.5%; Score 44; DB 2; Length 295;
Best Local Similarity 38.9%; Pred. No. 25;
Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 3 ASVKYVILDMYRALTLTM 20
:|:::|:::|:::|:::|

Db 13 SSIKVALLEIIIAAILVIM 30
:|:::|:::|:::|:::|

RESULT 5
S09882
hypothetical protein Ull15 - human cytomegalovirus (strain AD169)
C;Species: human cytomegalovirus, human herpesvirus 5
A;Note: host Homo sapiens (man)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C;Accession: S09882
R;Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;
M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A;Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A;Reference number: S09749; PMID:90269039; PMID:2161319
A;Accession: S09882
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-306 <CHE>
A;Cross-references: UNIPROT:P16932; EMBL:X17403; NID:g59591; PIDN:CAA35317.1; PID:e27311
A;Note: this sequence was submitted to the EMBL Data Library, December 1989
A;Note: this reading frame extends between two stop codons and does not begin with a start

Query Match 35.5%; Score 44; DB 2; Length 306;
Best Local Similarity 60.0%; Pred. No. 25;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 11 DMYRALTLMTSTA 25
|:::|:::|:::|:::|
Db 145 DQLRALTLSSDTA 159
|:::|:::|:::|:::|

RESULT 6
T39453
probable mrna stability protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T39453
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Churcher, C.M.
submitted to the EMBL Data Library, June 1998
A;Reference number: Z1855
A;Accession: T39453
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1242 <LYN>
A;Cross-references: UNIPROT:O60103; EMBL:AL023780; NID:el295812; PIDN:CAA19325.1; GSPDB:D
C;Genetics:
A;Experimental source: strain 972h-; cosmid cl4F5
A;Gene: SPDB:SPBC14F5.07
A;Map position: 2
A;Introns: 25/2; 38/2; 63/1

Query Match 35.5%; Score 44; DB 2; Length 1242;
Best Local Similarity 47.4%; Pred. No. 1.1e+02;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 ASVKYVILDMYRALTLTMN 21
|||:::|:::|:::|:::|
Db 1022 ASVKFLDYLRLSLNNKLN 1040
|||:::|:::|:::|:::|

RESULT 7
S69749
hypothetical protein YDR187c - Yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 23-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 29-Oct-1999
C;Accession: S69749
R;Murphy, L.; Harris, D.E.
submitted to the EMBL Data Library, November 1994
A;Reference number: S49764
A;Accession: S69749
A;Molecule type: DNA
A;Residues: 1-172 <MUR>

RESULT 12

D75104
transmembrane transport protein PAB1564 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: D75104
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome str
A:Reference number: A75001
A:Accession: D75104
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-503 <RAW>
A:Cross-references: UNIPROT:Q9UZC2; GB:AJ248286; GB:AL096836; NID:G5458366; PIDN:CAB5013
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB1564
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1141

Query Match 34.7%; Score 43; DB 2; Length 503;
Best Local Similarity 69.2%; Pred. No. 62;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EQASVKYVILDMY 13
| : |||||
Db 460 EKYGVKYLDPY 472

RESULT 13

T52187
probable transposase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C:Accession: T52187
R:Liu, D.; Crawford, N.M.
Genetics 149, 693-70, 1998
A:Title: Characterization of the putative transposase mRNA of Tag1, which is ubiquitous
A:Reference number: Z25992
A:Accession: T52187
A:Status: preliminary; translated from GB/EMBL/DBSJ
A:Molecule type: mRNA
A:Residues: 1-729 <LIU>
A:Cross-references: UNIPROT:O81190; EMBL:AF051562; PIDN:AAC25101.1

Query Match 34.7%; Score 43; DB 2; Length 729;
Best Local Similarity 47.6%; Pred. No. 90;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 EQASVKYVILDMYRALLTLN 21
| : |||||
Db 461 EKPSLPYVYEGMYRRLGIKN 481

RESULT 14

A55218
sfp homolog gsp - Bacillus brevis
C:Species: Bacillus brevis
C:Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 09-Jul-2004
C:Accession: A55218; A33593
R:Borchert, S.; Stachelhaus, T.; Marahiel, M.A.
J. Bacteriol. 176, 2458-2462, 1994
A:Title: Induction of surfactin production in Bacillus subtilis by gsp, a gene located u
A:Reference number: A55218; MUID:94209252; PMID:7512553
A:Accession: A55218
A:Molecule type: DNA
A:Residues: 1-237 <BOR>
A:Cross-references: UNIPROT:P40683; GB:X76434; NID:G474184; PIDN:CAA53988.1; PID:G474185
R:Kraetzschmar, J.; Krause, M.; Marahiel, M.A.
J. Bacteriol. 171, 5422-5429, 1989
A:Title: Gramicidin S biosynthesis operon containing the structural genes gsa and gsb

A:Reference number: A33593; MUID:90008776; PMID:2477357

A:Accession: A33593
A:Molecule type: DNA
A:Residues: 13-237 <KRA>
A:Cross-references: GB:X15577; NID:G39366; PIDN:CAA33601.1; PID:G39367
C:Genetics:
A:Gene: gsp
C:Superfamily: siderophore biosynthesis regulatory protein sfp

Query Match 34.3%; Score 42.5; DB 2; Length 237;
Best Local Similarity 55.6%; Pred. No. 34;
Matches 10; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 1 EQASVKYV-ILDMYRALL 17
| : |||||
Db 30 QQAFVRYVNVKDAYRSLL 47

RESULT 15

F75133
hypothetical protein PAB0581 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 16-Aug-2004
C:Accession: F75133
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome str
A:Reference number: A75001
A:Accession: F75133
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-474 <RAW>
A:Cross-references: UNIPROT:Q9V0C1; GB:AJ248285; GB:AL096836; NID:G5458067; PIDN:CAB497
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB0581
C:Superfamily: Peptidase V

Query Match 34.3%; Score 42.5; DB 2; Length 474;
Best Local Similarity 41.7%; Pred. No. 70;
Matches 10; Conservative 6; Mismatches 7; Indels 1; Gaps 1;

QY 1 EQASVKYVILDM-YRALLTLMNTS 23
| : |||||
Db 7 ELAKIQFLALDVQYRYLLDKLTKTS 30

Search completed: November 10, 2004, 14:52:23
Job time : 7.37736 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 : Search time 34.634 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-55

Sequence: 1 EQASVKVILDMYRALTLTMTSTAT 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	53	42.7	317 2	Q8RGL6
2	52	41.9	274 1	VGLL_MCMVK
3	52	41.9	274 1	VGLL_MCMVS
4	51	41.1	345 2	Q9ZUE4
5	51	41.1	375 2	Q9LQC7
6	51	41.1	428 2	O8RFT8
7	50	40.3	389 2	O97VN5
8	48	38.7	237 1	THX_BPPHC
9	48	38.7	357 2	Q7F007
10	48	38.7	357 2	Q9AUJ9
11	47.5	38.3	149 2	Q9K2L9
12	47	37.9	637 2	O6D999
13	47	37.9	680 2	Q7XPT9
14	46	37.1	262 2	O8SKT7
15	46	37.1	641 2	O8PST9
16	46	37.1	694 2	O8IK91
17	46	37.1	1065 2	Q7TDB4
18	45	36.3	158 2	Q8IIN4
19	45	36.3	270 2	Q7P4T7
20	45	36.3	299 2	Q38246
21	45	36.3	299 2	Q38292
22	45	36.3	661 2	Q6CMY7
23	45	36.3	884 2	Q9MA60
24	45	36.3	6885 1	SNE2_HUMAN
25	44.5	35.9	606 2	Q7KVA9
26	44.5	35.9	606 2	Aaf47526
27	44.5	35.9	709 2	Q7KVBI
28	44.5	35.9	709 2	Aaf47524
29	44.5	35.9	743 2	Q9W0C5
30	44.5	35.9	743 2	Aaf47525
31	44.5	35.9	1398 2	Q9XZ32

ALIGNMENTS

RESULT 1

Q8RGL6 PRELIMINARY; PRT; 317 AA.
AC O8RGL6
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Transposase.
GN OrderedLocusNames=FN0275;
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 25586;
RC MEDLINE=21886394; PubMed=11889109;
RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Fusch G., Haselkorn R.,
RA Fontstein M., Kyripides N.C., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018(2002).
DR EXML; AE010539; AAL94481.1;
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR009057; Homeodomain like.
DR InterPro; IPR002560; Transposase_12.
DR Pfam; PF01610; Transposase_12;
KW Complete proteome.
SQ SEQUENCE 317 AA; 37031 MW; 716DCAE60ABFC885 CRC64;

Query Match 42.7%; Score 53; DB 2; Length 317;

Best Local Similarity 42.1%; Pred. No. 3.5;

Matches 8; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 4 SVKXVILDMYRALTLTMT 22

DB 219 NVKYICMDMTYSISLVNS 237

RESULT 2

VGLL_MCMVK STANDARD; PRT; 274 AA.
ID VGLL_MCMVK
AC P52513;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DS Glycoprotein L precursor.
GN Name=GL; Synonyms=UL115;
OS Murine cytomegalovirus (strain KI81).

Q6C3P1 Yarrowia li
Q68667 human cytom
Q68668 human cytom
Q68669 human cytom
Q68670 human cytom
Q68671 human cytom
Q68672 human cytom
Q68673 human cytom
P16932 human cytom
Q68674 human cytom
Q68675 human cytom
Q71409 human cytom
Q71404 human cytom
Q7M6S9 human cytom

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Muromegalovirus.
 OX NCBI_TaxID=69156;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95053910; PubMed=7964634;
 RA Xu J., Scalzo A.A., Lyons P.A., Farrell H.E., Rawlinson W.D.,
 RA Shellam G.R.;
 RT "Identification, sequencing and expression of the glycoprotein L gene
 of murine cytomegalovirus.";
 RL J. Gen. Virol. 75:3235-3240(1994).
 CC -!- FUNCTION: Associated with glycoprotein H (gH) to form a complex
 CC important for infection and cell fusion. This association is
 CC necessary for the correct processing and cell surface expression
 CC of gH (By similarity).
 CC -!- SIMILARITY: Belongs to the herpesviruses glycoprotein L family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: L32962; AAA57343.1; -;
 DR InterPro: IPR002689; Cytomegalo_gL.
 DR Pfam: PF01801; Cytomegalo_gL; 1.
 KW Envelope protein; Glycoprotein; Signal.
 FT SIGNAL 1 17 Potential.
 FT CHAIN 18 274 Glycoprotein L.
 FT CARBOHYD 48 48 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 64 64 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 116 116 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 170 170 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 199 199 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 274 AA; 31239 MW; 504D349FDBEC5C03 CRC64;
 Query Match 41.9%; Score 52; DB 1; Length 274;
 Best Local Similarity 50.0%; Pred. No. 4.4;
 Matches 13; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
 QY 1 EQASVKYVILDMYRALLTLMTSTAT 26
 DB 109 EQSLHLNNTDQLRVLLTLMTSRAS 134
 RESULT 3
 ID VGLL_MCMVS STANDARD; PRT; 274 AA.
 AC P52514;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Glycoprotein L precursor.
 GN Name=GL; Synonyms=UL15;
 OS Murine cytomegalovirus (strain Smith).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Muromegalovirus.
 OX NCBI_TaxID=10367;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95053910; PubMed=7964634;
 RA Xu J., Scalzo A.A., Lyons P.A., Farrell H.E., Rawlinson W.D.,
 RA Shellam G.R.;
 RT "Identification, sequencing and expression of the glycoprotein L gene
 of murine cytomegalovirus.";
 RL J. Gen. Virol. 75:3235-3240(1994).
 CC -!- FUNCTION: Associated with glycoprotein H (gH) to form a complex
 CC important for infection and cell fusion. This association is
 CC necessary for the correct processing and cell surface expression
 CC of gH (By similarity).
 CC -!- SIMILARITY: Belongs to the herpesviruses glycoprotein L family.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: L32963; AAA57344.1; -;
 DR InterPro: IPR002689; Cytomegalo_gL.
 DR Pfam: PF01801; Cytomegalo_gL; 1.
 KW Envelope protein; Glycoprotein; Signal.
 FT SIGNAL 1 17 Potential.
 FT CHAIN 18 274 Glycoprotein L.
 FT CARBOHYD 48 48 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 64 64 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 116 116 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 170 170 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 199 199 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 274 AA; 31221 MW; 504D359EDBED5D03 CRC64;
 Query Match 41.9%; Score 52; DB 1; Length 274;
 Best Local Similarity 50.0%; Pred. No. 4.4;
 Matches 13; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
 QY 1 EQASVKYVILDMYRALLTLMTSTAT 26
 DB 109 EQSLHLNNTDQLRVLLTLMTSRAS 134
 RESULT 4
 ID Q9ZUE4 PRELIMINARY; PRT; 345 AA.
 AC Q9ZUE4;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE F508.6 protein.
 GN Name=F508.6;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Liu S., Lenz C.,
 RA Li J., Kremenetskaia I., Luros J., Altafi H., Gonzalez A., Araujo R.,
 RA Buehler E., Conn L., Conway A.B., Dunn P., Hansen N., Huizar L.,
 RA Kim C., Palm C.J., Rowley D., Shinn P., Walker M., Davis R.W.,
 RA Ecker J.R., Federspiel N.A., Theologis A.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Theologis;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC05990; AAC98006.1; -;
 DR PIR: F86368; F86368.
 DR GO: GO:0003824; F: catalytic activity; IEA.
 DR InterPro: IPR001087; Lipase_GDSL.
 DR InterPro: IPR008265; Lipase_GDSL_AS.
 DR Pfam: PF00657; Lipase_GDSL; 1.
 DR PROSITE: PS01098; LIPASE_GDSL_SER; 1.
 SQ SEQUENCE 345 AA; 37472 MW; 98368ABBB60A9C1E CRC64;
 Query Match 41.1%; Score 51; DB 2; Length 345;
 Best Local Similarity 47.6%; Pred. No. 8.3;
 Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
 QY 6 KYVILDMYRALLTLMTSTAT 26
 DB 271 KFYVDMYNPLNLLINNPRAS 291


```

RESULT 5
Q9LQC7 PRELIMINARY; PRT; 375 AA.
ID Q9LQC7
AC Q9LQC7
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F28C11.13.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Altafi H., Bei B., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
RA Ecker J.;
RA Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RL [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RA Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC007945; AAF79588.1; -.
DR GO; GO:0003824; P: catalytic activity; IEA.
DR InterPro; IPR001087; Lipase_GDSL.
DR Pfam; PF00657; Lipase_GDSL; 1.
DR PROSITE; PS01098; Lipase_GDSL_SER; 1.
SQ SEQUENCE 375 AA; 41143 MW; F1EFA4D13BB7243 CRC64;

Query Match 41.1%; Score 51; DB 2; Length 375;
Best Local Similarity 47.6%; Pred. No. 9.1;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 6 KXVILDMYRALTLNMTSTAT 26
Db 301 KFYVDMYNPLINPRAS 321

RESULT 6
Q8RFT8 PRELIMINARY; PRT; 428 AA.
ID Q8RFT8
AC Q8RFT8
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transposase.
GN OrderedLocusNames=FN0599;
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;

```

```

OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fongstein M., Kyprides N.C., Overbeek R.;
RA "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586."
J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AE010571; AAL94795.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR009057; Homeodomain-like.
DR InterPro; IPR002560; Transposase_12.
DR Pfam; PF01610; Transposase_12; 1.
KW Complete proteome.
SQ SEQUENCE 428 AA; 49937 MW; BA99FA4348FAB3C3 CRC64;

Query Match 41.1%; Score 51; DB 2; Length 428;
Best Local Similarity 42.1%; Pred. No. 10;
Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 4 SVKYVILDMYRALTLNMT 22
Db 219 NVKICMDMTPYISLVNS 237

RESULT 7
Q97VNS PRELIMINARY; PRT; 389 AA.
ID Q97VNS
AC Q97VNS;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sulfolipid biosynthesis protein (Sgdb).
GN Namesgdb: OrderedLocusNames=SSO2583;
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Brauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Koza C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;
RA "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
RT Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
RL EMBL; AE006855; AAK42708.1; -.
DR PIR; E90431; E90431.
DR HSSP; O48917; 1QRR.
KW Complete proteome.
SQ SEQUENCE 389 AA; 44686 MW; 1A63ADFF33A26C3 CRC64;

Query Match 40.3%; Score 50; DB 2; Length 389;
Best Local Similarity 42.3%; Pred. No. 14;
Matches 11; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 EQASVKYVILDMYRALTLNMTSTAT 26
Db 99 EQRSAPYSMDVNHVNTLNNLTST 124

```

```

RESULT 8
THYX BPPHC
ID THYX BPPHC STANDARD; PRT; 237 AA.
AC Q2X32;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 08-JUL-2004 (Rel. 44, Last annotation update)
DE Probable thymidylate synthase (EC 2.1.1.148) (TS) (TSase) (GP16).
GN Name=16;
OS Bacteriophage phi-C31.
OS Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=10719;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Norwich;
RX MEDLINE=99162580; PubMed=10051617;
RA Hendrix R.W., Smith M.C.M., Burns N., Ford M.E., Hatfull G.F.;
RT "Evolutionary relationships among diverse bacteriophages and
RT prophages: all the world's a phage.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:2192-2197(1999).
CC -!- FUNCTION: Catalyzes the formation of dTMP and tetrahydrofolate
CC from dUMP and methylenetetrahydrofolate (By similarity).
CC -!- CATALYTIC ACTIVITY: 5,10-methylenetetrahydrofolate + dUMP +
CC FADH(2) = dTMP + tetrahydrofolate + FAD.
CC -!- COFACTOR: Binds 1 FAD per subunit (By similarity).
CC -!- SIMILARITY: Belongs to the thymidylate synthase thyX family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ006589; CA007140.1; -
CC InterPro; IPR003669; Thyl.
CC Pfam; PF02511; Thyl; 1.
CC PIR; P02511; Thyl; 1.
CC K1; K1; Thyl; 1.
CC K2; K2; Thyl; 1.
CC K3; K3; Thyl; 1.
CC K4; K4; Thyl; 1.
CC K5; K5; Thyl; 1.
CC K6; K6; Thyl; 1.
CC K7; K7; Thyl; 1.
CC K8; K8; Thyl; 1.
CC K9; K9; Thyl; 1.
CC K10; K10; Thyl; 1.
CC K11; K11; Thyl; 1.
CC K12; K12; Thyl; 1.
CC K13; K13; Thyl; 1.
CC K14; K14; Thyl; 1.
CC K15; K15; Thyl; 1.
CC K16; K16; Thyl; 1.
CC K17; K17; Thyl; 1.
CC K18; K18; Thyl; 1.
CC K19; K19; Thyl; 1.
CC K20; K20; Thyl; 1.
CC K21; K21; Thyl; 1.
CC K22; K22; Thyl; 1.
CC K23; K23; Thyl; 1.
CC K24; K24; Thyl; 1.
CC K25; K25; Thyl; 1.
CC K26; K26; Thyl; 1.
CC K27; K27; Thyl; 1.
CC K28; K28; Thyl; 1.
CC K29; K29; Thyl; 1.
CC K30; K30; Thyl; 1.
CC K31; K31; Thyl; 1.
CC K32; K32; Thyl; 1.
CC K33; K33; Thyl; 1.
CC K34; K34; Thyl; 1.
CC K35; K35; Thyl; 1.
CC K36; K36; Thyl; 1.
CC K37; K37; Thyl; 1.
CC K38; K38; Thyl; 1.
CC K39; K39; Thyl; 1.
CC K40; K40; Thyl; 1.
CC K41; K41; Thyl; 1.
CC K42; K42; Thyl; 1.
CC K43; K43; Thyl; 1.
CC K44; K44; Thyl; 1.
CC K45; K45; Thyl; 1.
CC K46; K46; Thyl; 1.
CC K47; K47; Thyl; 1.
CC K48; K48; Thyl; 1.
CC K49; K49; Thyl; 1.
CC K50; K50; Thyl; 1.
CC K51; K51; Thyl; 1.
CC K52; K52; Thyl; 1.
CC K53; K53; Thyl; 1.
CC K54; K54; Thyl; 1.
CC K55; K55; Thyl; 1.
CC K56; K56; Thyl; 1.
CC K57; K57; Thyl; 1.
CC K58; K58; Thyl; 1.
CC K59; K59; Thyl; 1.
CC K60; K60; Thyl; 1.
CC K61; K61; Thyl; 1.
CC K62; K62; Thyl; 1.
CC K63; K63; Thyl; 1.
CC K64; K64; Thyl; 1.
CC K65; K65; Thyl; 1.
CC K66; K66; Thyl; 1.
CC K67; K67; Thyl; 1.
CC K68; K68; Thyl; 1.
CC K69; K69; Thyl; 1.
CC K70; K70; Thyl; 1.
CC K71; K71; Thyl; 1.
CC K72; K72; Thyl; 1.
CC K73; K73; Thyl; 1.
CC K74; K74; Thyl; 1.
CC K75; K75; Thyl; 1.
CC K76; K76; Thyl; 1.
CC K77; K77; Thyl; 1.
CC K78; K78; Thyl; 1.
CC K79; K79; Thyl; 1.
CC K80; K80; Thyl; 1.
CC K81; K81; Thyl; 1.
CC K82; K82; Thyl; 1.
CC K83; K83; Thyl; 1.
CC K84; K84; Thyl; 1.
CC K85; K85; Thyl; 1.
CC K86; K86; Thyl; 1.
CC K87; K87; Thyl; 1.
CC K88; K88; Thyl; 1.
CC K89; K89; Thyl; 1.
CC K90; K90; Thyl; 1.
CC K91; K91; Thyl; 1.
CC K92; K92; Thyl; 1.
CC K93; K93; Thyl; 1.
CC K94; K94; Thyl; 1.
CC K95; K95; Thyl; 1.
CC K96; K96; Thyl; 1.
CC K97; K97; Thyl; 1.
CC K98; K98; Thyl; 1.
CC K99; K99; Thyl; 1.
CC K100; K100; Thyl; 1.
CC K101; K101; Thyl; 1.
CC K102; K102; Thyl; 1.
CC K103; K103; Thyl; 1.
CC K104; K104; Thyl; 1.
CC K105; K105; Thyl; 1.
CC K106; K106; Thyl; 1.
CC K107; K107; Thyl; 1.
CC K108; K108; Thyl; 1.
CC K109; K109; Thyl; 1.
CC K110; K110; Thyl; 1.
CC K111; K111; Thyl; 1.
CC K112; K112; Thyl; 1.
CC K113; K113; Thyl; 1.
CC K114; K114; Thyl; 1.
CC K115; K115; Thyl; 1.
CC K116; K116; Thyl; 1.
CC K117; K117; Thyl; 1.
CC K118; K118; Thyl; 1.
CC K119; K119; Thyl; 1.
CC K120; K120; Thyl; 1.
CC K121; K121; Thyl; 1.
CC K122; K122; Thyl; 1.
CC K123; K123; Thyl; 1.
CC K124; K124; Thyl; 1.
CC K125; K125; Thyl; 1.
CC K126; K126; Thyl; 1.
CC K127; K127; Thyl; 1.
CC K128; K128; Thyl; 1.
CC K129; K129; Thyl; 1.
CC K130; K130; Thyl; 1.
CC K131; K131; Thyl; 1.
CC K132; K132; Thyl; 1.
CC K133; K133; Thyl; 1.
CC K134; K134; Thyl; 1.
CC K135; K135; Thyl; 1.
CC K136; K136; Thyl; 1.
CC K137; K137; Thyl; 1.
CC K138; K138; Thyl; 1.
CC K139; K139; Thyl; 1.
CC K140; K140; Thyl; 1.
CC K141; K141; Thyl; 1.
CC K142; K142; Thyl; 1.
CC K143; K143; Thyl; 1.
CC K144; K144; Thyl; 1.
CC K145; K145; Thyl; 1.
CC K146; K146; Thyl; 1.
CC K147; K147; Thyl; 1.
CC K148; K148; Thyl; 1.
CC K149; K149; Thyl; 1.
CC K150; K150; Thyl; 1.
CC K151; K151; Thyl; 1.
CC K152; K152; Thyl; 1.
CC K153; K153; Thyl; 1.
CC K154; K154; Thyl; 1.
CC K155; K155; Thyl; 1.
CC K156; K156; Thyl; 1.
CC K157; K157; Thyl; 1.
CC K158; K158; Thyl; 1.
CC K159; K159; Thyl; 1.
CC K160; K160; Thyl; 1.
CC K161; K161; Thyl; 1.
CC K162; K162; Thyl; 1.
CC K163; K163; Thyl; 1.
CC K164; K164; Thyl; 1.
CC K165; K165; Thyl; 1.
CC K166; K166; Thyl; 1.
CC K167; K167; Thyl; 1.
CC K168; K168; Thyl; 1.
CC K169; K169; Thyl; 1.
CC K170; K170; Thyl; 1.
CC K171; K171; Thyl; 1.
CC K172; K172; Thyl; 1.
CC K173; K173; Thyl; 1.
CC K174; K174; Thyl; 1.
CC K175; K175; Thyl; 1.
CC K176; K176; Thyl; 1.
CC K177; K177; Thyl; 1.
CC K178; K178; Thyl; 1.
CC K179; K179; Thyl; 1.
CC K180; K180; Thyl; 1.
CC K181; K181; Thyl; 1.
CC K182; K182; Thyl; 1.
CC K183; K183; Thyl; 1.
CC K184; K184; Thyl; 1.
CC K185; K185; Thyl; 1.
CC K186; K186; Thyl; 1.
CC K187; K187; Thyl; 1.
CC K188; K188; Thyl; 1.
CC K189; K189; Thyl; 1.
CC K190; K190; Thyl; 1.
CC K191; K191; Thyl; 1.
CC K192; K192; Thyl; 1.
CC K193; K193; Thyl; 1.
CC K194; K194; Thyl; 1.
CC K195; K195; Thyl; 1.
CC K196; K196; Thyl; 1.
CC K197; K197; Thyl; 1.
CC K198; K198; Thyl; 1.
CC K199; K199; Thyl; 1.
CC K200; K200; Thyl; 1.
CC K201; K201; Thyl; 1.
CC K202; K202; Thyl; 1.
CC K203; K203; Thyl; 1.
CC K204; K204; Thyl; 1.
CC K205; K205; Thyl; 1.
CC K206; K206; Thyl; 1.
CC K207; K207; Thyl; 1.
CC K208; K208; Thyl; 1.
CC K209; K209; Thyl; 1.
CC K210; K210; Thyl; 1.
CC K211; K211; Thyl; 1.
CC K212; K212; Thyl; 1.
CC K213; K213; Thyl; 1.
CC K214; K214; Thyl; 1.
CC K215; K215; Thyl; 1.
CC K216; K216; Thyl; 1.
CC K217; K217; Thyl; 1.
CC K218; K218; Thyl; 1.
CC K219; K219; Thyl; 1.
CC K220; K220; Thyl; 1.
CC K221; K221; Thyl; 1.
CC K222; K222; Thyl; 1.
CC K223; K223; Thyl; 1.
CC K224; K224; Thyl; 1.
CC K225; K225; Thyl; 1.
CC K226; K226; Thyl; 1.
CC K227; K227; Thyl; 1.
CC K228; K228; Thyl; 1.
CC K229; K229; Thyl; 1.
CC K230; K230; Thyl; 1.
CC K231; K231; Thyl; 1.
CC K232; K232; Thyl; 1.
CC K233; K233; Thyl; 1.
CC K234; K234; Thyl; 1.
CC K235; K235; Thyl; 1.
CC K236; K236; Thyl; 1.
CC K237; K237; Thyl; 1.
CC K238; K238; Thyl; 1.
CC K239; K239; Thyl; 1.
CC K240; K240; Thyl; 1.
CC K241; K241; Thyl; 1.
CC K242; K242; Thyl; 1.
CC K243; K243; Thyl; 1.
CC K244; K244; Thyl; 1.
CC K245; K245; Thyl; 1.
CC K246; K246; Thyl; 1.
CC K247; K247; Thyl; 1.
CC K248; K248; Thyl; 1.
CC K249; K249; Thyl; 1.
CC K250; K250; Thyl; 1.
CC K251; K251; Thyl; 1.
CC K252; K252; Thyl; 1.
CC K253; K253; Thyl; 1.
CC K254; K254; Thyl; 1.
CC K255; K255; Thyl; 1.
CC K256; K256; Thyl; 1.
CC K257; K257; Thyl; 1.
CC K258; K258; Thyl; 1.
CC K259; K259; Thyl; 1.
CC K260; K260; Thyl; 1.
CC K261; K261; Thyl; 1.
CC K262; K262; Thyl; 1.
CC K263; K263; Thyl; 1.
CC K264; K264; Thyl; 1.
CC K265; K265; Thyl; 1.
CC K266; K266; Thyl; 1.
CC K267; K267; Thyl; 1.
CC K268; K268; Thyl; 1.
CC K269; K269; Thyl; 1.
CC K270; K270; Thyl; 1.
CC K271; K271; Thyl; 1.
CC K272; K272; Thyl; 1.
CC K273; K273; Thyl; 1.
CC K274; K274; Thyl; 1.
CC K275; K275; Thyl; 1.
CC K276; K276; Thyl; 1.
CC K277; K277; Thyl; 1.
CC K278; K278; Thyl; 1.
CC K279; K279; Thyl; 1.
CC K280; K280; Thyl; 1.
CC K281; K281; Thyl; 1.
CC K282; K282; Thyl; 1.
CC K283; K283; Thyl; 1.
CC K284; K284; Thyl; 1.
CC K285; K285; Thyl; 1.
CC K286; K286; Thyl; 1.
CC K287; K287; Thyl; 1.
CC K288; K288; Thyl; 1.
CC K289; K289; Thyl; 1.
CC K290; K290; Thyl; 1.
CC K291; K291; Thyl; 1.
CC K292; K292; Thyl; 1.
CC K293; K293; Thyl; 1.
CC K294; K294; Thyl; 1.
CC K295; K295; Thyl; 1.
CC K296; K296; Thyl; 1.
CC K297; K297; Thyl; 1.
CC K298; K298; Thyl; 1.
CC K299; K299; Thyl; 1.
CC K300; K300; Thyl; 1.
CC K301; K301; Thyl; 1.
CC K302; K302; Thyl; 1.
CC K303; K303; Thyl; 1.
CC K304; K304; Thyl; 1.
CC K305; K305; Thyl; 1.
CC K306; K306; Thyl; 1.
CC K307; K307; Thyl; 1.
CC K308; K308; Thyl; 1.
CC K309; K309; Thyl; 1.
CC K310; K310; Thyl; 1.
CC K311; K311; Thyl; 1.
CC K312; K312; Thyl; 1.
CC K313; K313; Thyl; 1.
CC K314; K314; Thyl; 1.
CC K315; K315; Thyl; 1.
CC K316; K316; Thyl; 1.
CC K317; K317; Thyl; 1.
CC K318; K318; Thyl; 1.
CC K319; K319; Thyl; 1.
CC K320; K320; Thyl; 1.
CC K321; K321; Thyl; 1.
CC K322; K322; Thyl; 1.
CC K323; K323; Thyl; 1.
CC K324; K324; Thyl; 1.
CC K325; K325; Thyl; 1.
CC K326; K326; Thyl; 1.
CC K327; K327; Thyl; 1.
CC K328; K328; Thyl; 1.
CC K329; K329; Thyl; 1.
CC K330; K330; Thyl; 1.
CC K331; K331; Thyl; 1.
CC K332; K332; Thyl; 1.
CC K333; K333; Thyl; 1.
CC K334; K334; Thyl; 1.
CC K335; K335; Thyl; 1.
CC K336; K336; Thyl; 1.
CC K337; K337; Thyl; 1.
CC K338; K338; Thyl; 1.
CC K339; K339; Thyl; 1.
CC K340; K340; Thyl; 1.
CC K341; K341; Thyl; 1.
CC K342; K342; Thyl; 1.
CC K343; K343; Thyl; 1.
CC K344; K344; Thyl; 1.
CC K345; K345; Thyl; 1.
CC K346; K346; Thyl; 1.
CC K347; K347; Thyl; 1.
CC K348; K348; Thyl; 1.
CC K349; K349; Thyl; 1.
CC K350; K350; Thyl; 1.
CC K351; K351; Thyl; 1.
CC K352; K352; Thyl; 1.
CC K353; K353; Thyl; 1.
CC K354; K354; Thyl; 1.
CC K355; K355; Thyl; 1.
CC K356; K356; Thyl; 1.
CC K357; K357; Thyl; 1.
CC K358; K358; Thyl; 1.
CC K359; K359; Thyl; 1.
CC K360; K360; Thyl; 1.
CC K361; K361; Thyl; 1.
CC K362; K362; Thyl; 1.
CC K363; K363; Thyl; 1.
CC K364; K364; Thyl; 1.
CC K365; K365; Thyl; 1.
CC K366; K366; Thyl; 1.
CC K367; K367; Thyl; 1.
CC K368; K368; Thyl; 1.
CC K369; K369; Thyl; 1.
CC K370; K370; Thyl; 1.
CC K371; K371; Thyl; 1.
CC K372; K372; Thyl; 1.
CC K373; K373; Thyl; 1.
CC K374; K374; Thyl; 1.
CC K375; K375; Thyl; 1.
CC K376; K376; Thyl; 1.
CC K377; K377; Thyl; 1.
CC K378; K378; Thyl; 1.
CC K379; K379; Thyl; 1.
CC K380; K380; Thyl; 1.
CC K381; K381; Thyl; 1.
CC K382; K382; Thyl; 1.
CC K383; K383; Thyl; 1.
CC K384; K384; Thyl; 1.
CC K385; K385; Thyl; 1.
CC K386; K386; Thyl; 1.
CC K387; K387; Thyl; 1.
CC K388; K388; Thyl; 1.
CC K389; K389; Thyl; 1.
CC K390; K390; Thyl; 1.
CC K391; K391; Thyl; 1.
CC K392; K392; Thyl; 1.
CC K393; K393; Thyl; 1.
CC K394; K394; Thyl; 1.
CC K395; K395; Thyl; 1.
CC K396; K396; Thyl; 1.
CC K397; K397; Thyl; 1.
CC K398; K398; Thyl; 1.
CC K399; K399; Thyl; 1.
CC K400; K400; Thyl; 1.
CC K401; K401; Thyl; 1.
CC K402; K402; Thyl; 1.
CC K403; K403; Thyl; 1.
CC K404; K404; Thyl; 1.
CC K405; K405; Thyl; 1.
CC K406; K406; Thyl; 1.
CC K407; K407; Thyl; 1.
CC K408; K408; Thyl; 1.
CC K409; K409; Thyl; 1.
CC K410; K410; Thyl; 1.
CC K411; K411; Thyl; 1.
CC K412; K412; Thyl; 1.
CC K413; K413; Thyl; 1.
CC K414; K414; Thyl; 1.
CC K415; K415; Thyl; 1.
CC K416; K416; Thyl; 1.
CC K417; K417; Thyl; 1.
CC K418; K418; Thyl; 1.
CC K419; K419; Thyl; 1.
CC K420; K420; Thyl; 1.
CC K421; K421; Thyl; 1.
CC K422; K422; Thyl; 1.
CC K423; K423; Thyl; 1.
CC K424; K424; Thyl; 1.
CC K425; K425; Thyl; 1.
CC K426; K426; Thyl; 1.
CC K427; K427; Thyl; 1.
CC K428; K428; Thyl; 1.
CC K429; K429; Thyl; 1.
CC K430; K430; Thyl; 1.
CC K431; K431; Thyl; 1.
CC K432; K432; Thyl; 1.
CC K433; K433; Thyl; 1.
CC K434; K434; Thyl; 1.
CC K435; K435; Thyl; 1.
CC K436; K436; Thyl; 1.
CC K437; K437; Thyl; 1.
CC K438; K438; Thyl; 1.
CC K439; K439; Thyl; 1.
CC K440; K440; Thyl; 1.
CC K441; K441; Thyl; 1.
CC K442; K442; Thyl; 1.
CC K443; K443; Thyl; 1.
CC K444; K444; Thyl; 1.
CC K445; K445; Thyl; 1.
CC K446; K446; Thyl; 1.
CC K447; K447; Thyl; 1.
CC K448; K448; Thyl; 1.
CC K449; K449; Thyl; 1.
CC K450; K450; Thyl; 1.
CC K451; K451; Thyl; 1.
CC K452; K452; Thyl; 1.
CC K453; K453; Thyl; 1.
CC K454; K454; Thyl; 1.
CC K455; K455; Thyl; 1.
CC K456; K456; Thyl; 1.
CC K457; K457; Thyl; 1.
CC K458; K458; Thyl; 1.
CC K459; K459; Thyl; 1.
CC K460; K460; Thyl; 1.
CC K461; K461; Thyl; 1.
CC K462; K462; Thyl; 1.
CC K463; K463; Thyl; 1.
CC K464; K464; Thyl; 1.
CC K465; K465; Thyl; 1.
CC K466; K466; Thyl; 1.
CC K467; K467; Thyl; 1.
CC K468; K468; Thyl; 1.
CC K469; K469; Thyl; 1.
CC K470; K470; Thyl; 1.
CC K471; K471; Thyl; 1.
CC K472; K472; Thyl; 1.
CC K473; K473; Thyl; 1.
CC K474; K474; Thyl; 1.
CC K475; K475; Thyl; 1.
CC K476; K476; Thyl; 1.
CC K477; K477; Thyl; 1.
CC K478; K478; Thyl; 1.
CC K479; K479; Thyl; 1.
CC K480; K480; Thyl; 1.
CC K481; K481; Thyl; 1.
CC K482; K482; Thyl; 1.
CC K483; K483; Thyl; 1.
CC K484; K484; Thyl; 1.
CC K485; K485; Thyl; 1.
CC K486; K486; Thyl; 1.
CC K487; K487; Thyl; 1.
CC K488; K488; Thyl; 1.
CC K489; K489; Thyl; 1.
CC K490; K490; Thyl; 1.
CC K491; K491; Thyl; 1.
CC K492; K492; Thyl; 1.
CC K493; K493; Thyl; 1.
CC K494; K494; Thyl; 1.
CC K495; K495; Thyl; 1.
CC K496; K496; Thyl; 1.
CC K497; K497; Thyl; 1.
CC K498; K498; Thyl; 1.
CC K499; K499; Thyl; 1.
CC K500; K500; Thyl; 1.
CC K501; K501; Thyl; 1.
CC K502; K502; Thyl; 1.
CC K503; K503; Thyl; 1.
CC K504; K504; Thyl; 1.
CC K505; K505; Thyl; 1.
CC K506; K506; Thyl; 1.
CC K507; K507; Thyl; 1.
CC K508; K508; Thyl; 1.
CC K509; K509; Thyl; 1.
CC K510; K510; Thyl; 1.
CC K511; K511; Thyl; 1.
CC K512; K512; Thyl; 1.
CC K513; K513; Thyl; 1.
CC K514; K514; Thyl; 1.
CC K515; K515; Thyl; 1.
CC K516; K516; Thyl; 1.
CC K517; K517; Thyl; 1.
CC K518; K518; Thyl; 1.
CC K519; K519; Thyl; 1.
CC K520; K520; Thyl; 1.
CC K521; K521; Thyl; 1.
CC K522; K522; Thyl; 1.
CC K523; K523; Thyl; 1.
CC K524; K524; Thyl; 1.
CC K525; K525; Thyl; 1.
CC K526; K526; Thyl; 1.
CC K527; K527; Thyl; 1.
CC K528; K528; Thyl; 1.
CC K529; K529; Thyl; 1.
CC K530; K530; Thyl; 1.
CC K531; K531; Thyl; 1.
CC K532; K532; Thyl; 1.
CC K533; K533; Thyl; 1.
CC K534; K534; Thyl; 1.
CC K535; K535; Thyl; 1.
CC K536; K536; Thyl; 1.
CC K537; K537; Thyl; 1.
CC K538; K538; Thyl; 1.
CC K539; K539; Thyl; 1.
CC K540; K540; Thyl; 1.
CC K541; K541; Thyl; 1.
CC K542; K542; Thyl; 1.
CC K543; K543; Thyl; 1.
CC K544; K544; Thyl; 1.
CC K545; K545; Thyl; 1.
CC K546; K546; Thyl; 1.
CC K547; K547; Thyl; 1.
CC K548; K548; Thyl; 1.
CC K549; K549; Thyl; 1.
CC K550; K550; Thyl; 1.
CC K551; K551; Thyl; 1.
CC K552; K552; Thyl; 1.
CC K553; K553; Thyl; 1.
CC K554; K554; Thyl; 1.
CC K555; K555; Thyl; 1.
CC K556; K556; Thyl; 1.
CC K557; K557; Thyl; 1.
CC K558; K558; Thyl; 1.
CC K559; K559; Thyl; 1.
CC K560; K560; Thyl; 1.
CC K561; K561; Thyl; 1.
CC K562; K562; Thyl; 1.
CC K563; K563; Thyl; 1.
CC K564; K564; Thyl; 1.
CC K565; K565; Thyl; 1.
CC K566; K566; Thyl; 1.
CC K567; K567; Thyl; 1.
CC K568; K568; Thyl; 1.
CC K569; K569; Thyl; 1.
CC K570; K570; Thyl; 1.
CC K571; K571; Thyl; 1.
CC K572; K572; Thyl; 1.
CC K573; K573; Thyl; 1.
CC K574; K574; Thyl; 1.
CC K575; K575; Thyl; 1.
CC K576; K576; Thyl; 1.
CC K577; K577; Thyl; 1.
CC K578; K578; Thyl; 1.
CC K579; K579; Thyl; 1.
CC K580; K580; Thyl; 1.
CC K581; K581; Thyl; 1.
CC K582; K582; Thyl; 1.
CC K583; K583; Thyl; 1.
CC K584; K584; Thyl; 1.
CC K585; K585; Thyl; 1.
CC K586; K586; Thyl; 1.
CC K587; K587; Thyl; 1.
CC K588; K588; Thyl; 1.
CC K589; K589; Thyl; 1.
CC K590; K590; Thyl; 1.
CC K591; K591; Thyl; 1.
CC K592; K592; Thyl; 1.
CC K593; K593; Thyl; 1.
CC K594; K594; Thyl; 1.
CC K595; K595; Thyl; 1.
CC K596; K596; Thyl; 1.
CC K597; K597; Thyl; 1.
CC K598; K598; Thyl; 1.
CC K599; K599; Thyl; 1.
CC K600; K600; Thyl; 1.
CC K601; K601; Thyl; 1.
CC K602; K602; Thyl; 1.
CC K603; K603; Thyl; 1.
CC K604; K604; Thyl; 1.
CC K605; K605; Thyl; 1.
CC K606; K606; Thyl; 1.
CC K607; K607; Thyl; 1.
CC K608; K608; Thyl; 1.
CC K609; K609; Thyl; 1.
CC K610; K610; Thyl; 1.
CC K611; K611; Thyl; 1.
CC K612; K612; Thyl; 1.
CC K613; K613; Thyl; 1.
CC K614; K614; Thyl; 1.
CC K615; K615; Thyl; 1.
CC K616; K616; Thyl; 1.
CC K617; K617; Thyl; 1.
CC K618; K618; Thyl; 1.
CC K619; K619; Thyl; 1.
CC K620; K620; Thyl; 1.
CC K621; K621; Thyl; 1.
CC K622; K622; Thyl; 1.
CC K623; K623; Thyl; 1.
CC K624; K624; Thyl; 1.
CC K625; K625; Thyl; 1.
CC K626; K626; Thyl; 1.
CC K627; K627; Thyl; 1.
CC K628; K628; Thyl; 1.
CC K629; K629; Thyl; 1.
CC K630; K630; Thyl; 1.
CC K631; K631; Thyl; 1.
CC K632; K632; Thyl; 1.
CC K633; K633; Thyl; 1.
CC K634; K634; Thyl; 1.
CC K635; K635; Thyl; 1.
CC K636; K636; Thyl; 1.
CC K637; K637; Thyl; 1.
CC K638; K638; Thyl; 1.
CC K639; K639; Thyl; 1.
CC K640; K640; Thyl; 1.
CC K641; K641; Thyl; 1.
CC K642; K642; Thyl; 1.
CC K643; K643; Thyl; 1.
CC K644; K644; Thyl; 1.
CC K645; K645; Thyl; 1.
CC K646; K646; Thyl; 1.
CC K647; K647; Thyl; 1.
CC K648; K648; Thyl; 1.
CC K649; K649; Thyl; 1.
CC K650; K650; Thyl; 1.
CC K651; K651; Thyl; 1.
CC K652; K652; Thyl; 1.
CC K653; K653; Thyl; 1.
CC K654; K654; Thyl; 1.
```

```
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (Fragment).
GN Name=gapDH;
OS Erwinia tracheiphila.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=65700;
RN [1]_TaxID=65700;
RP SEQUENCE FROM N.A.
RC STRAIN=ET9815, and ET5845;
RX MEDLINE=21027675; PubMed=11155980;
RA Brown E.W., Davis R.M., Gouk C., van der Zwet T.;
RT "Phylogenetic relationships of necrogenic Erwinia and Brenneria
RT species as revealed by glyceraldehyde-3-phosphate dehydrogenase gene
RT sequences.";
RL Int. J. Syst. Evol. Microbiol. 50:2057-2068 (2000).
CC -!- SIMILARITY: Belongs to the glyceraldehyde-3-phosphate
CC dehydrogenase family.
DR EMBL; AF165012; AAF31217.1; -.
DR EMBL; AF165011; AAF31216.1; -.
DR HSSP; P06977; 1DC5.
DR GO; GO:0004365; P:glyceraldehyde-3-phosphate dehydrogenase (p. .: IEA.
DR GO; GO:006491; P:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000173; GAP_dhhdhrogenase.
DR Pfam; PF02800; Gp dh C; 1.
DR Pfam; PF00044; Gp dh N; 1.
DR PRINTS; PR00078; G3PDHDSGNASE.
KW Oxidoreductase.
FT NON_TER 1
FT NON_TER 149
FT SEQUENCE 149 AA; 16376 MW; D89D17E8858E0122 CRC64;
SQ
Query Match 38.3%; Score 47.5; DB 2; Length 149;
Best Local Similarity 44.4%; Pred. No. 13;
Matches 12; Conservative 7; Mismatches 5; Indels 3; Gaps 1;
QY 3 ASVKYVILDMY---RALTLMTSTAT 26
DB 104 APLAYVINDKFGIVGALITVVHTTAT 130
RESULT 12
Q6D999 PRELIMINARY; PRT; 637 AA.
AC Q6D999;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE ATP-independent RNA helicase.
GN Name=dead; Synonyms=msb; ORFNames=ECA0718;
OS Erwinia carotovora subsp. atroseptica SCR11043.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=218491;
RN [1]_TaxID=218491;
RP SEQUENCE FROM N.A.
RC STRAIN=SCR11043;
RA Bell K.S., Sebainia M., Pritchard L., Holden M., Hyman L.J.,
RA Holveva M.C., Thomson N.R., Bentley S.D., Churcher C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Frazer A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmon G.P.C., Birch P.R.J., Barrall B.G., Parkhill J., Toth I.K.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; EX950851; CAG73633.1; -.
KW Helicase.
SQ SEQUENCE 637 AA; 70916 MW; 7CAC606D8B216B07 CRC64;
Query Match 37.9%; Score 47; DB 2; Length 637;
Best Local Similarity 50.9%; Pred. No. 76;
```

```
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 10 LDMYRALLTLM 20
DB 400 LDMYRALLTLM 410
RESULT 13
Q7XPT9 PRELIMINARY; PRT; 680 AA.
AC Q7XPT9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE OSJNBa0088H09.17 protein.
GN Names=OSJNBa0088H09.17;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L., Li T.,
RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
RA Zhang Y., Hu H., Jia P.X., Qian Y.W., Ying K., Zhou B., Chen Z.H.,
RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu X.F., Tu Y.F., Jia J., Yin H.F.,
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL606651; CAE03459.1; -.
DR Gramene; Q7XPT9; -.
DR GO; GO:0000145; C:exocyst; IEA.
DR GO; GO:0006887; P:exocytosis; IEA.
DR InterPro; IPR001917; Aminotrans_I.
DR InterPro; IPR004140; Exo70.
DR Pfam; PF03081; Exo70; 1.
DR PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN 1.
SQ SEQUENCE 680 AA; 77228 MW; 86A5D7F09F7464B9 CRC64;
Query Match 37.9%; Score 47; DB 2; Length 680;
Best Local Similarity 53.3%; Pred. No. 81;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 7 YVILDMYRALLTLMN 21
DB 339 YILDMYKATLELQS 353
RESULT 14
Q86KT7 PRELIMINARY; PRT; 262 AA.
AC Q86KT7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA MEDLINE=22092622; PubMed=12097910;
RA Gloeckner G., Eichinger L., Szafarski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RL "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RT Nature 418:79-85 (2002).
RN [2]
```

Search completed: November 10, 2004, 14:50:20
Job time : 36.634 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 7.32075 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-56

Perfect score: 103
Sequence: 1 EDLESVLRINWAKGSP1 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A_COMB.psp: *
2: /cgn2_6/ptodata/1/iaa/5B_COMB.psp: *
3: /cgn2_6/ptodata/1/iaa/6A_COMB.psp: *
4: /cgn2_6/ptodata/1/iaa/6B_COMB.psp: *
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.psp: *
6: /cgn2_6/ptodata/1/iaa/backfiles1.psp: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	44.7	176	US-09-602-472A-61	Sequence 61, Appl
2	46	44.7	652	US-08-261-663A-6	Sequence 6, Appl
3	46	44.7	652	US-09-357-206A-5	Sequence 5, Appl
4	46	44.7	652	US-09-813-742A-5	Sequence 5, Appl
5	46	44.7	652	PCT-US95-07754A-6	Sequence 6, Appl
6	46	44.7	1143	US-08-310-912A-108	Sequence 108, App
7	46	44.7	1143	US-09-301-085-108	Sequence 108, App
8	46	44.7	1143	PCT-US95-04589-108	Sequence 108, App
9	46	44.7	1144	US-08-261-663A-2	Sequence 2, Appl
10	46	44.7	1144	US-08-261-663A-4	Sequence 4, Appl
11	46	44.7	1144	US-08-930-996A-9	Sequence 9, Appl
12	46	44.7	1144	US-09-357-206A-3	Sequence 3, Appl
13	46	44.7	1144	US-09-813-742A-3	Sequence 3, Appl
14	46	44.7	1144	PCT-US95-07754A-2	Sequence 2, Appl
15	46	44.7	1144	PCT-US95-07754A-4	Sequence 4, Appl
16	45	43.7	317	US-08-605-284B-17	Sequence 17, Appl
17	45	43.7	378	US-09-710-279-1428	Sequence 1428, Ap
18	45	43.7	387	US-09-134-001C-3248	Sequence 3248, Ap
19	44	42.7	119	US-08-467-420A-19	Sequence 19, Appl
20	44	42.7	119	US-08-470-110A-19	Sequence 19, Appl
21	44	42.7	119	US-08-667-769A-19	Sequence 19, Appl
22	44	42.7	119	US-08-940-371-19	Sequence 19, Appl
23	44	42.7	119	US-08-637-647-19	Sequence 19, Appl
24	44	42.7	119	PCT-US95-17082A-19	Sequence 19, Appl
25	44	42.7	393	US-09-107-532A-3712	Sequence 3712, Ap
26	44	42.7	1551	US-09-437-568A-46	Sequence 46, Appl
27	44	42.7	2780	US-10-220-587-2	Sequence 2, Appl

Sequence 78, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 68, Appl
Sequence 66, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 8, Appl
Sequence 9, Appl
Sequence 10, Appl
Sequence 1, Appl
Patent No. 523606
Sequence 17, Appl
Sequence 35, Appl
Sequence 12, Appl
Sequence 12, Appl

28 43 41.7 131 4 US-09-631-594-78
29 43 41.7 284 4 US-09-387-375-7
30 43 41.7 284 4 US-10-041-400A-7
31 43 41.7 316 4 US-09-387-375-9
32 43 41.7 316 4 US-10-041-400A-9
33 43 41.7 513 4 US-09-071-035-68
34 43 41.7 532 4 US-09-071-035-66
35 43 41.7 600 4 US-09-134-000C-6286
36 42 40.8 97 4 US-09-652-345-5
37 42 40.8 243 2 US-08-836-620A-8
38 42 40.8 243 2 US-08-836-620A-9
39 42 40.8 243 2 US-08-836-620A-10
40 42 40.8 264 4 US-09-560-979-1
41 42 40.8 410 6 5223608-5
42 42 40.8 591 2 US-08-836-620A-17
43 42 40.8 593 3 US-09-041-886-35
44 42 40.8 593 3 US-08-764-870-12
45 42 40.8 593 3 US-08-980-115-12

ALIGNMENTS

RESULT 1

US-09-602-472A-61
; Sequence 61, Application US/09602472A
; Patent No. 6608240
; GENERAL INFORMATION:
; APPLICANT: Bidney, Dennis L.
; APPLICANT: Hu, Xu
; APPLICANT: Li, Guihua
; TITLE OF INVENTION: Sunflower Disease Resistance Genes
; FILE REFERENCE: 35718/200630
; CURRENT APPLICATION NUMBER: US/09/602,472A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/140,876
; PRIOR FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 61
; TYPE: PRT
; LENGTH: 176
; ORGANISM: Nicotiana tabacum
US-09-602-472A-61

Query Match 44.7% Score 46; DB 4; Length 176;
Best Local Similarity 36.8%; Pred. No. 9.4;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EDLESVLRINWAKGSP1 19
Db 154 ENFEKLSLVVNVAKGLPL 172

RESULT 2

US-08-261-663A-6
; Sequence 6, Application US/08261663A
; Patent No. 5571706
; GENERAL INFORMATION:
; APPLICANT: Baker, Barbara J
; APPLICANT: Whitham, Steven A
; TITLE OF INVENTION: Plant Virus Resistance Gene and Methods
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Margaret A. Connor, USDA-ARS
; STREET: 800 Buchanan Street
; CITY: Albany
; STATE: CA
; COUNTRY: USA
; ZIP: 94710
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/261,663A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Connor, Margaret A.
; REGISTRATION NUMBER: 30043
; REFERENCE/DOCKET NUMBER: 0094.94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 559-6067
; TELEFAX: (510) 559-5777
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 652 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-261-663A-6

Query Match 44.7%; Score 46; DB 1; Length 652;
Best Local Similarity 36.8%; Pred. No. 41;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EDLESVLRLINWAKGSP1 19
Db 369 ENFEKLSLEVVNYAKGLPL 387

RESULT 3
US-09-357-206A-5
; Sequence 5, Application US/09357206A
; Patent No. 6372962
; GENERAL INFORMATION:
; APPLICANT: Dinesh-Kumar, S.
; APPLICANT: Baker, Barbara
; TITLE OF INVENTION: Pathogen Resistance in Plants using cDNA-N/Intron Constructs
; FILE REFERENCE: 042250/191805 (5830-5)
; CURRENT APPLICATION NUMBER: US/09/357,206A
; CURRENT FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: US 60/093,494
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 5
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Nicotiana glutinosa
US-09-357-206A-5

Query Match 44.7%; Score 46; DB 3; Length 652;
Best Local Similarity 36.8%; Pred. No. 41;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EDLESVLRLINWAKGSP1 19
Db 369 ENFEKLSLEVVNYAKGLPL 387

RESULT 4
US-09-813-742A-5
; Sequence 5, Application US/09813742A
; Patent No. 6630618
; GENERAL INFORMATION:
; APPLICANT: Baker, Barbara J
; APPLICANT: Dinesh-Kumar, S.P.
; TITLE OF INVENTION: NON-PATHOGEN INDUCED SYSTEMIC ACQUIRED RESISTANCE (SAR) IN PLANTS
; FILE REFERENCE: 42250/209601 (5830-12)
; CURRENT APPLICATION NUMBER: US/09/813,742A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,027
; PRIOR FILING DATE: 2000-03-21
```

```
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Nicotiana glutinosa
US-09-813-742A-5

Query Match 44.7%; Score 46; DB 4; Length 652;
Best Local Similarity 36.8%; Pred. No. 41;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EDLESVLRLINWAKGSP1 19
Db 369 ENFEKLSLEVVNYAKGLPL 387

RESULT 5
PCT-US95-07754A-6
; Sequence 6, Application PC/TUS9507754A
; GENERAL INFORMATION:
; APPLICANT: Baker, Barbara J
; APPLICANT: Whitham, Steven A
; TITLE OF INVENTION: Plant Virus Resistance Gene and Methods
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Margaret A. Connor, USDA-ARS
; STREET: 800 Buchanan Street
; CITY: Albany
; STATE: CA
; COUNTRY: USA
; ZIP: 94710
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07754A
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Connor, Margaret A.
; REGISTRATION NUMBER: 30043
; REFERENCE/DOCKET NUMBER: 0094.94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 559-6067
; TELEFAX: (510) 559-5777
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 652 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-07754A-6

Query Match 44.7%; Score 46; DB 5; Length 652;
Best Local Similarity 36.8%; Pred. No. 41;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EDLESVLRLINWAKGSP1 19
Db 369 ENFEKLSLEVVNYAKGLPL 387

RESULT 6
US-08-310-912A-108
; Sequence 108, Application US/08310912A
; Patent No. 5981730
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Staskawicz, Brian J.
; APPLICANT: Brent, Andrew F.
```

APPLICANT: Dahlbeck, Douglas
APPLICANT: Katagiri, Fumiaki
APPLICANT: Kunkel, Barbara N.
APPLICANT: Mindrinos, Michael N.
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND DETECTION
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 208
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2904
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/310,912A
FILING DATE: September 22, 1994
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/227,360
FILING DATE: April 13, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Lech, Karen F.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/254001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 100254
INFORMATION FOR SEQ ID NO: 108:
SEQUENCE CHARACTERISTICS:
LENGTH: 1143 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-310-912A-108

Query Match 44.7%; Score 46; DB 2; Length 1143;
Best Local Similarity 36.8%; Pred. No. 77;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 EDLESVLRLINWAKGSP1 19
DB 369 ENFEKLSLEVNYAKGLPL 387

RESULT 7
US-09-301-085-108
Sequence 108, Application US/09301085
GENERAL INFORMATION:
APPLICANT: Ausubel, Frederick M.
APPLICANT: Staskawicz, Brian J.
APPLICANT: Brent, Andrew F.
APPLICANT: Dahlbeck, Douglas
APPLICANT: Katagiri, Fumiaki
APPLICANT: Kunkel, Barbara N.
APPLICANT: Mindrinos, Michael N.
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND
FILE REFERENCE: 00786/254002
CURRENT APPLICATION NUMBER: US/09/301,085
CURRENT FILING DATE: 1999-04-28
EARLIER APPLICATION NUMBER: 08/310,912
EARLIER FILING DATE: 1994-09-22
EARLIER APPLICATION NUMBER: 08/227,360

EARLIER FILING DATE: 1994-04-13
NUMBER OF SEQ ID NOS: 208
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 108
LENGTH: 1143
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-301-085-108

Query Match 44.7%; Score 46; DB 3; Length 1143;
Best Local Similarity 36.8%; Pred. No. 77;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 EDLESVLRLINWAKGSP1 19
DB 369 ENFEKLSLEVNYAKGLPL 387

RESULT 8
PCT-US95-04589-108
Sequence 108, Application PC/TUS9504589
GENERAL INFORMATION:
APPLICANT: Ausubel, Frederick M.
APPLICANT: Staskawicz, Brian J.
APPLICANT: Brent, Andrew F.
APPLICANT: Dahlbeck, Douglas
APPLICANT: Katagiri, Fumiaki
APPLICANT: Kunkel, Barbara N.
APPLICANT: Mindrinos, Michael N.
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: RPS2 GENE AND USES THEREOF
NUMBER OF SEQUENCES: 201
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2904
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04589
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,360
FILING DATE: 13-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/230001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 100254
INFORMATION FOR SEQ ID NO: 108:
SEQUENCE CHARACTERISTICS:
LENGTH: 1143 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-04589-108

Query Match 44.7%; Score 46; DB 5; Length 1143;
Best Local Similarity 36.8%; Pred. No. 77;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 EDLESVLRLINWAKGSP1 19
DB 369 ENFEKLSLEVNYAKGLPL 387


```

; APPLICANT: Whitham, Steven A
; TITLE OF INVENTION: Plant Virus Resistance Gene and Methods
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Margaret A. Connor, USDA-ARS
; STREET: 800 Buchanan Street
; CITY: Albany
; STATE: CA
; COUNTRY: USA
; ZIP: 94710
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07754A
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Connor, Margaret A
; REGISTRATION NUMBER: 30043
; REFERENCE/DOCKET NUMBER: 0094.94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 559-6067
; TELEFAX: (510) 559-5777
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1144 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-07754-A-2

Query Match          44.7%; Score 46; DB 5; Length 1144;
Best Local Similarity 36.8%; Pred. No. 77;
Matches 7; Conservative 7; Mismatches 5; Indels

QY      1 EDLESLIRLINWAKGSP I 19
       |:|:::|:|:|:|:|
Db      369 ENFEKLSLEVANYAKGLPL 387

RESULT 15
PCT-US95-07754-A-4
; Sequence 4, Application PC/TUS9507754A
; GENERAL INFORMATION:
; APPLICANT: Baker, Barbara J
; APPLICANT: Whitham, Steven A
; TITLE OF INVENTION: Plant Virus Resistance Gene and Methods
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Margaret A. Connor, USDA-ARS
; STREET: 800 Buchanan Street
; CITY: Albany
; STATE: CA
; COUNTRY: USA
; ZIP: 94710
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07754A
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Connor, Margaret A
; REGISTRATION NUMBER: 30043
; REFERENCE/DOCKET NUMBER: 0094.94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 559-6067

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 23.1321 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-56

Perfect score: 103
Sequence: 1 EDLESVLRLINWAKGSPIP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	103	100.0	20	14	US-10-092-750-56
2	50	48.5	226	16	US-10-767-701-44097
3	49	47.6	367	16	US-10-437-963-175166
4	49	47.6	393	16	US-10-437-963-175162
5	49	47.6	1288	14	US-10-114-824A-10
6	48	46.6	778	9	US-09-789-919-50
7	47	45.6	103	15	US-10-424-599-211799
8	47	45.6	108	15	US-10-424-599-265311
9	46	44.7	217	15	US-10-424-599-163276
10	46	44.7	423	14	US-10-363-493-8926
11	46	44.7	445	17	US-10-425-115-194216
12	46	44.7	652	9	US-09-813-742-5
13	46	44.7	1143	9	US-09-867-852-108

14	44.7	1143	15	US-10-613-472-108	Sequence 108, Appl
15	44.7	1143	16	US-10-613-765-108	Sequence 108, Appl
16	44.7	1144	9	US-09-813-742-3	Sequence 3, Appl
17	45	43.7	72	US-10-424-599-212155	Sequence 212155,
18	45	43.7	798	US-10-161-572-51	Sequence 51, Appl
19	45	43.7	1163	US-10-295-027-1256	Sequence 153840,
20	44.5	43.2	49	US-10-437-963-153840	Sequence 68106, A
21	44.5	43.2	425	US-10-425-114-68106	Sequence 281104,
22	44.5	43.2	439	US-10-424-599-281104	Sequence 164301,
23	44.5	43.2	1503	US-10-437-963-164301	Sequence 164299,
24	44.5	43.2	1524	US-10-437-963-164299	Sequence 138100,
25	44	42.7	84	US-10-437-963-138100	Sequence 19, Appl
26	44	42.7	119	US-10-144-644-19	Sequence 19, Appl
27	44	42.7	119	US-10-700-740-19	Sequence 285321,
28	44	42.7	125	US-10-425-115-285321	Sequence 215229,
29	44	42.7	254	US-10-424-599-215229	Sequence 59795, A
30	44	42.7	311	US-10-425-114-59795	Sequence 59796, A
31	44	42.7	386	US-10-425-115-337442	Sequence 337442,
32	44	42.7	401	US-10-425-115-337442	Sequence 2361, Ap
33	44	42.7	428	US-10-389-566-2361	Sequence 47027, A
34	44	42.7	430	US-10-282-122A-47207	Sequence 70149, A
35	44	42.7	833	US-10-282-122A-70149	Sequence 46, Appl
36	44	42.7	833	US-10-318-906A-46	Sequence 46, Appl
37	44	42.7	1551	US-10-319-236A-46	Sequence 12, Appl
38	44	42.7	1551	US-10-418-036-12	Sequence 2, Appl
39	44	42.7	1551	US-10-618-039-12	Sequence 2, Appl
40	44	42.7	1551	US-10-220-587-2	Sequence 2, Appl
41	44	42.7	2780	US-10-423-483-2	Sequence 218873,
42	44	42.7	2780	US-10-424-599-218873	Sequence 31646, A
43	43.5	42.2	150	US-10-029-386-31646	Sequence 7, Appl
44	43	41.7	49	US-09-832-197-7	
45	43	41.7	52	US-09-832-197-7	

ALIGNMENTS

RESULT 1
US-10-092-750-56
; Sequence 56, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-56

Query Match 100.0%; Score 103; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDLESVLRLINWAKGSPIP 20
Db 1 EDLESVLRLINWAKGSPIP 20

RESULT 2
US-10-767-701-44097
; Sequence 44097, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:


```

; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 778
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: All Xaa's in this sequence represent unreadable
; OTHER INFORMATION: amino acid.
US-09-789-919-50

Query Match          46.6%; Score 48; DB 9; Length 778;
Best Local Similarity 47.1%; Pred.No. 94;
Matches            8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      2 DLSVLRLINWKGSP 18
      ||::||:::||
DB      472 DLQFLINIVFNCKNSP 488

RESULT 7
US-10-424-599-211799
; Sequence 211799, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 211799
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_33281C.1.pep
US-10-424-599-211799

Query Match          45.6%; Score 47; DB 15; Length 103;
Best Local Similarity 70.0%; Pred.No. 15;
Matches            7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      9 RLINWKGSP 18
      ||::||:::||
DB      84 RVLNWKGP 93

RESULT 8
US-10-424-599-265311
; Sequence 265311, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 265311
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT MRT3847 81596C.1.pep

```

```
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Staskawicz, Brian J.
; APPLICANT: Brent, Andrew F.
; APPLICANT: Dahlbeck, Douglas
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kunkel, Barbara N.
; APPLICANT: Mindrinos, Michael N.
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND
; DETECTION METHODS
; FILE REFERENCE: 00786/254002
; CURRENT APPLICATION NUMBER: US/09/867,852
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/301,085
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/310,912
; PRIOR FILING DATE: EARLIER FILING DATE: 1994-09-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/227,360
; PRIOR FILING DATE: EARLIER FILING DATE: 1994-04-13
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 108
; LENGTH: 1143
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-867-852-108

Query Match          44.7%; Score 46; DB 9; Length 1143;
Best Local Similarity 36.8%; Pred.No. 3ex+02;
Matches      7; Conservative    7; Mismatches     5; Indels   0;

QY      1 EDLESVLRLINWAKGSP I 19
         | : : : : : ||| | :
Db       369 ENFEKLSLEVNVYAKGLPL 387

RESULT 14
US-10-613-472-108
; Sequence 108, Application US/10613472
; Publication No. US20040088756A1
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Staskawicz, Brian J.
; APPLICANT: Brent, Andrew F.
; APPLICANT: Dahlbeck, Douglas
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kunkel, Barbara N.
; APPLICANT: Mindrinos, Michael N.
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND
; DETECTION METHODS
; FILE REFERENCE: 00786/254004
; CURRENT APPLICATION NUMBER: US/10/613,472
; CURRENT FILING DATE: 2003-07-02
; PRIOR APPLICATION NUMBER: US 09/867,852
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 09/301,085
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: US 08/310,912
; PRIOR FILING DATE: 1994-09-22
; PRIOR APPLICATION NUMBER: US 08/227,360
; PRIOR FILING DATE: 1994-04-13
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 108
; LENGTH: 1143
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-613-472-108

Query Match          44.7%; Score 46; DB 15; Length 1143;
Best Local Similarity 36.8%; Pred.No. 3ex+02;
```

Matches	7;	Conservative	7;	Mismatches	5;	Indels	0;	Gaps	0;
---------	----	--------------	----	------------	----	--------	----	------	----

```
QY      1 EDLESVLIRLINWAKGSP I 19
         || : | : : ||| | :
Db     369 ENFEKLSLEV VNYAKGLPL 387
```

RESULT 15

```

US-10-613-765-108
; Sequence 108, Application US/10613765
; Publication No. US20040172673A1
; GENERAL INFORMATION:
; APPLICANT: AUSUBEL, FREDERICK M.
; APPLICANT: STASKAWICZ, BRIAN J.
; APPLICANT: KATAGIRI, FUMIAKI
; APPLICANT: KUNKEL, BARBARA N.
; APPLICANT: MINDRINOS, MICHAEL N.
; APPLICANT: YU, GUO-LIANG
; APPLICANT: BAKER, BARBARA
; APPLICANT: ELLIS, JEFFREY
; APPLICANT: SALMERON, JOHN
; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, AND
; TITLE OF INVENTION: DETECTION METHODS
; FILE REFERENCE: 00786/254005
; CURRENT APPLICATION NUMBER: US/10/613,765
; CURRENT FILING DATE: 2003-07-02
; PRIOR APPLICATION NUMBER: US 09/867,852
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 09/301,085
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: US 08/310,912
; PRIOR FILING DATE: 1994-09-22
; PRIOR APPLICATION NUMBER: US 08/227,360
; PRIOR FILING DATE: 1994-04-13
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 108
; LENGTH: 1143
; TYPE: FRT
; ORGANISM: Arabidopsis thaliana
US-10-613-765-108

```

```
QY      1 EDLESVLIRLINWAXGSP1 19
        | : | : | : | : | : | :
Dp     369 ENFEKLSLEVNYAKGLPL 387
```

Search completed: November 11, 2004, 02:43:12
Job time : 24.1821 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run On: November 10, 2004, 13:40:53 ; Search time 4.90566 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-56

Perfect score: 103
Sequence: 1 EDLESVLRLINWAKGSPIP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	48.5	392	B32370	cyclin B2 - Africa
2	49	47.6	935	E96806	hypothetical prote
3	48	46.6	20	S59382	hypothetical prote
4	48	46.6	1647	T41267	hypothetical prote
5	47	45.6	848	JG0194	androgen receptor
6	46	44.7	756	T48139	copper amine oxida
7	46	44.7	1144	A54810	TMV resistance pro
8	46	44.7	1186	H88869	protein unc-31 (im
9	45	43.7	250	E72605	probable high-affi
10	45	43.7	297	F70201	conserved hypotet
11	45	43.7	381	D89922	conserved hypotet
12	45	43.7	397	A32370	cyclin B1 - Africa
13	45	43.7	431	T29267	hypothetical prote
14	45	43.7	611	T41563	hypothetical prote
15	45	43.7	1522	JC1101	sodium channel pro
16	44.5	43.2	170	T08821	disease resistance
17	44.5	43.2	326	AF2024	hypothetical prote
18	44.5	43.2	414	A72765	hypothetical prote
19	44.5	43.2	583	S19476	hypothetical prote
20	44	42.7	244	1 LNRTMC	mannose-binding le
21	44	42.7	290	A23519	isochorismatase (E
22	44	42.7	394	T42512	ID-myo-inositol-tr
23	44	42.7	429	AD3642	(S)-2-hydroxy-acid
24	44	42.7	430	C70176	probable ATP-depen
25	44	42.7	461	T42513	ID-myo-inositol-tr
26	44	42.7	486	T25639	hypothetical prote
27	44	42.7	494	T42444	ID-myo-inositol-tr
28	44	42.7	548	1 DJSPS2	DNA-directed DNA p
29	44	42.7	725	T35114	probable Kinase/pi

MRP-like ABC trans
conserved hypotet
hypothetical prote
hypothetical prote
cycloin B - rat
cycloin B - long-ta
cycloin B1 - mouse
cycloin B - mouse
SMV2 protein - Yea
cobalamin biosynth
glutamate-ammonia-
hypothetical prote
mannose-binding le
lactose transport
hypothetical prote

ALIGNMENTS

RESULT 1

B32370

cyclin B2 - African clawed frog

C;Species: Xenopus laevis (African clawed frog)

C;Date: 02-Nov-1989 #sequence_revision 02-Nov-1989 #text_change 12-Jul-2004

C;Accession: B32370

R;Minshull, J.; Blow, J.J.; Hunt, T.

Cell 56, 947-956, 1989

A;Title: Translation of cyclin mRNA is necessary for extracts of activated Xenopus eggs

A;Reference number: A32370; MUID:89168446; PMID:2564315

A;Accession: B32370

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-392 <MIN>

A;Cross-references: UNIPROT:P13351; GB:J03167; NID:G214094; PIDN:AAA49697.1; PID:G21409

C;Superfamily: cyclin, A/B/D/E type

C;Keywords: cell cycle control

Query Match

Best Local Similarity 48.5%; Score 50; DB 2; Length 392;

Matches 7; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Qy 2 DLESVLRLINWAKGSPIP 20

Db 245 EMEMILRLNFDLGRPLP 263

RESULT 2

E96806

hypothetical protein T32E8.1 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: E96806

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: E96806

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-935 <SFO>

A;Cross-references: UNIPROT:Q9CA26; GB:AF005173; NID:96437530; PIDN:AAF08562.1; GSPDB:G

C;Genetics: T32E8.1

A;Gene: T32E8.1

A;Map position: 1

Db 369 ENFEKLSLEVNYAKGLPL 387

RESULT 8

H8869

protein unc-31 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: H8869

R:anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; MUID:99069613; PMID:9851916

A>Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.ele

A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: H8869

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1186 <STO>

A:Cross-references: UNIPROT:Q23658; GB:chr_IV; PIDN:CAA93520.1; PID:g3881885; GSPDB:GN00

C:Genetics:

A:Gene: unc-31

A:Map position: 4

C:Superfamily: calcium-dependent actin-binding protein; pleckstrin repeat homology

Query Match 44.7%; Score 46; DB 2; Length 1186;

Best Local Similarity 58.8%; Pred. No. 56;

Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 3 LESVLRLINWAKGSP1 19

Db 1154 LESVLSRLARYDEGNPI 1170

RESULT 9

E72605

probable high-affinity branched-chain amino acid transport ATP-binding protein APE1308 -

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 16-Aug-2004

C:Accession: E72605

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: E72605

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-250 <KAW>

A:Cross-references: UNIPROT:Q9YCE9; DDBJ:AP000061; NID:g5104821; PIDN:BAA80299.1; PID:dl

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1308

C:Superfamily: ATP-binding cassette homology

Query Match 43.7%; Score 45; DB 2; Length 250;

Best Local Similarity 46.7%; Pred. No. 16;

Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LESVLRLINWAKGS 17

Db 94 LENINVRHPWTRGS 108

RESULT 10

F70201

conserved hypothetical protein BB0815 - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004

C:Accession: F70201

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White

son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,

; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A>Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: F70201

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-297 <KLE>

A:Cross-references: UNIPROT:O51755; GB:AE001180; GB:AE000783; NID:g2698755; PIDN:AAC671

A:Experimental source: strain B31

C:Superfamily: conserved hypothetical protein HI0176

Query Match 43.7%; Score 45; DB 2; Length 297;

Best Local Similarity 47.1%; Pred. No. 19;

Matches 8; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LESVLRLINWAKGSP1 19

Db 20 LDSILIKILNFKASII 36

RESULT 11

D89922

conserved hypothetical protein SA1277 [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: D89922

R:Kuroda, M.; Ohca, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A>Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: D89922

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-381 <KUR>

A:Cross-references: UNIPROT:Q99U44; GB:BA000018; PID:g13701242; PIDN:BAB42537.1; GSPDB:

A:Experimental source: strain N315

C:Genetics:

A:Gene: SA1277

C:Superfamily: Bacillus conserved hypothetical protein ypsc

Query Match 43.7%; Score 45; DB 2; Length 381;

Best Local Similarity 47.4%; Pred. No. 25;

Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 EDLESVLRLINWAKGSP1 19

Db 174 ETLAASLIRLANWKGDTPL 192

RESULT 12

A32370

cyclin B1 - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C>Date: 02-Nov-1989 #sequence_revision 02-Nov-1989 #text_change 12-Jul-2004

C:Accession: A32370

R:Minghull, J.; Blow, J.J.; Hunt, T.

Cell 56, 947-956, 1989

A>Title: translation of cyclin mRNA is necessary for extracts of activated Xenopus eggs

A:Reference number: A32370; MUID:89168446; PMID:2564315

A:Accession: A32370

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-397 <MIN>

A:Cross-references: UNIPROT:P13350; GB:J03166; NID:g214092; PIDN:AAA49696.1; PID:g21409

C:Superfamily: cyclin, A/B/D/E type

C:Keywords: cell cycle control

Query Match 43.7%; Score 45; DB 2; Length 397;

Best Local Similarity 36.8%; Pred. No. 26;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:18:57 ; Search time 26.6415 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-56

Perfect score: 103
Sequence: 1 EDLSVLRLINWAKGSPIP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : UniProt 02: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	49.5	471	2 Q88GL4	P88g14 pseudomonas
2	50	48.5	392	1 QGB2_XENLA	P13351 xenopus lae
3	50	48.5	761	1 QD24_MOUSE	Q9r160 mus musculu
4	50	48.5	761	2 Q8CDV3	Q8cdv3 mus musculu
5	49	47.6	133	2 Q7ND57	Q7nd57 Gloeobacter
6	49	47.6	935	2 Q9CA26	Q9ca26 arabidopsis
7	49	47.6	1187	2 Q8GZ83	Q8gz83 arabidopsis
8	49	47.6	1187	2 Q9FH84	Q9fh84 arabidopsis
9	48.5	47.1	150	2 Q84Y05	Q84y05 glycine max
10	48	46.6	594	2 Q745W1	Q745w1 thermus the
11	48	46.6	594	2 AA82524	AA82524 thermus t
12	48	46.6	760	2 Q8X1N6	Q8x1n6 cephalospor
13	48	46.6	1104	2 Q922R7	Q922r7 rattus norv
14	48	46.6	1647	2 P78847	P78847 schizosacch
15	47.5	46.1	554	2 Q9SF12	Q9sf12 arabidopsis
16	47	45.6	137	2 Q81H29	Q81h29 bacillus ce
17	47	45.6	147	2 Q8UWF7	Q8uwf7 salmo salar
18	47	45.6	166	2 Q9DDA0	Q9dda0 xenopus tro
19	47	45.6	392	2 Q6FA39	Q6fa39 xenopus lae
20	47	45.6	392	2 AAH60466	AAH60466 xenopus l
21	47	45.6	743	2 Q97120	Q97120 schistosoma
22	47	45.6	771	2 Q800S7	Q800s7 acanthopagr
23	47	45.6	848	2 Q9YGV9	Q9ygv9 anguilla ja
24	47	45.6	853	2 Q93245	Q93245 oncorhynch
25	46	44.7	128	2 Q9PUL1	Q9pull cryzias lat
26	46	44.7	130	2 Q9OWA3	Q9owa3 spheeroides
27	46	44.7	200	2 Q918F5	Q918f5 pinophales
28	46	44.7	310	2 Q73109	Q73109 wolbachia p
29	46	44.7	310	2 AA514103	AA514103 wolbachia
30	46	44.7	362	2 Q9T9H3	Q9t9h3 halocynthia
31	46	44.7	377	2 Q6EAL5	Q6eal5 brachydanio

32 46 44.7 563 2 Q9DDJ4
33 46 44.7 638 2 BAD12595
34 46 44.7 682 2 Q801Z1
35 46 44.7 735 2 Q801Z2
36 46 44.7 741 2 Q8L742
37 46 44.7 744 2 Q76LMS
38 46 44.7 744 2 BAC98301
39 46 44.7 756 2 Q9STI2
40 46 44.7 763 2 Q8QFV7
41 46 44.7 763 2 Q8UWB7
42 46 44.7 769 2 Q93497
43 46 44.7 839 2 Q8QFV2
44 46 44.7 854 2 Q93244
45 46 44.7 1002 2 Q7R290

ALIGNMENTS

RESULT 1

Q88GL4 PRELIMINARY; PRT; 471 AA.
AC Q88GL4;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocuNames=PP3707;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP MEDLINE=22423060; PubMed=12534463;
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Moazzaz A., Utterback T.R., Rizzo M., Lee K., Kosack D., Woelfel D.,
RA Wedler H., Lauber J., Stjepandic D., Hohnselt J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duisterhoef A., Tuemmler B.,
RA Fraser C.M.;
RA "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AEO16788; AAN69304.1; -.
DR TIGR; PP3707; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 471 AA; 52925 MW; 20A9B3CE128BDEF5 CRC64;

Query Match 49.5%; Score 51; DB 2; Length 471;
Best Local Similarity 42.1%; Pred. No. 16;
Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EDLSVLRLINWAKGSP 19

Db 449 EAMEKLLIKVQDMLGSPV 467

RESULT 2

QGB2_XENLA STANDARD; PRT; 392 AA.
AC P13351;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE G2/mitotic-specific cyclin B2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.

OX NCBI_TaxID=8355;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=89168446; PubMed=2564315;
 RX Minshull J., Blow J.J., Hunt T.;
 RA "Translation of cyclin mRNA is necessary for extracts of activated
 RT xenopus eggs to enter mitosis.";
 RL Cell 56:947-956(1989).
 CC -1- FUNCTION: Essential for the control of the cell cycle at the G2/M
 CC (mitosis) transition.
 CC -1- SUBUNIT: Interacts with the CDC2 protein kinase to form a
 CC serine/threonine kinase holoenzyme complex also known as
 CC maturation promoting factor (MPF). The cyclin subunit imparts
 CC substrate specificity to the complex.
 CC -1- DEVELOPMENTAL STAGE: Accumulates steadily during G2 and is
 CC abruptly destroyed at mitosis.
 CC -1- SIMILARITY: Belongs to the cyclin family. Cyclin AB subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; J03167; AAA49697.1; -.
 DR PIR; B32370; B32370.
 DR HSSP; P20248; 1H1R.
 DR InterPro; IPR006670; Cyclin.
 DR InterPro; IPR004367; Cyclin_Cterm.
 DR InterPro; IPR011028; Cyclin_Like.
 DR InterPro; IPR006671; Cyclin_N.
 DR Pfam; PF02984; Cyclin_C; 1.
 DR Pfam; PF00134; Cyclin_N; 1.
 DR SMART; SM00385; CYCLIN; 2.
 DR PROSITE; PS00292; CYCLINS; 1.
 KW Cell cycle; Cell division; Cyclin; Mitosis.
 SQ SEQUENCE 392 AA; 43624 MW; 87955E04C106218E CRC64;

 Query Match 48.5%; Score 50; DB 1; Length 392;
 Best Local Similarity 36.8%; Pred. No. 19;
 Matches 7; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

 QY 2 DLESLVRLINWAKGSPIP 20
 Db 245 EMEMILRLNFDLGRPLP 263

 RESULT 3
 AD24 MOUSE STANDARD; PRT; 761 AA.
 ID AD24 MOUSE
 AC Q9RI60;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE ADAM 24 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
 DE domain 24) (Testase 1).
 GN Name=Adam24;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Testis;
 RC MEDLINE=99326496; PubMed=10395895;
 RX Zhu G.-Z., Lin Y., Myles D.G., Primakoff P.;
 RA "Identification of four novel ADAMs with potential roles in
 RT spermatogenesis and fertilization.";
 RL Gene 234:227-237(1999).
 RN [2]
 RP CHARACTERIZATION.

RX MEDLINE=21206048; PubMed=11309208;
 RA Zhu G.-Z., Myles D.G., Primakoff P.;
 RT "Testase 1 (ADAM 24) a plasma membrane-anchored sperm protease
 RT implicated in sperm function during epididymal maturation or
 RL fertilization.";
 RL J. Cell Sci. 114:1787-1794(2001).
 CC -1- FUNCTION: Plasma membrane protease present on mature sperm that
 CC may be involved in sperm function during epididymal maturation
 CC and/or fertilization.
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -1- SUBUNIT: Monomer.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Expressed exclusively in testis and more
 CC specifically on the surface of mature sperm. Localized to the
 CC equatorial region of the plasma membrane of cauda epididymal
 CC sperm.
 CC -1- DEVELOPMENTAL STAGE: Adult levels are reached by day 20 after
 CC birth.
 CC -1- PM: The prodomain is removed during sperm passage through the
 CC caput epididymis after the protein has reached the cell surface.
 CC Not processed in the secretory pathway.
 CC -1- SIMILARITY: Belongs to peptidase family M12B.
 CC -1- SIMILARITY: Contains 1 disintegrin domain.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF167402; AAD48841.1; -.
 DR HSP; P18619; IFVL.
 DR MEROPS; M12.227; -.
 DR MGD; MGI:105984; Adam24.
 DR InterPro; IPR006586; ADAM cysteine.
 DR InterPro; IPR001762; Disintegrin.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR006209; EGF_Like.
 DR InterPro; IPR001590; Peptidase_M12B.
 DR InterPro; IPR002870; Peptidase_M12B_N.
 DR InterPro; IPR001818; Pept_M10A_M12B.
 DR InterPro; IPR006025; Pept_M_Zn_BS.
 DR Pfam; PF00200; Disintegrin; 1.
 DR Pfam; PF01562; Pep_M12B_propep; 1.
 DR Pfam; PF01421; Reprolysin; 1.
 DR ProDom; PD000664; Disintegrin; 1.
 DR SMART; SM00608; ACR; 1.
 DR SMART; SM00050; DISIN; 1.
 DR PROSITE; PS50215; ADAM_MEROPS; 1.
 DR PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
 DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
 DR PROSITE; PS00214; DISINTEGRIN_2; 1.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS50026; EGF_3; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW EGF-like domain; Glycoprotein; Hydrolase; Metalloprotease; Signal;
 KW Spermatogenesis; Transmembrane; Zinc; Zymogen.
 FT SIGNAL 35 34 Potential.
 FT PROPEP 1 200 By similarity.
 FT CHAIN 201 761 ADAM 24.
 FT DOMAIN 35 697 Extracellular (Potential).
 FT TRANSMEM 698 718 Potential.
 FT DOMAIN 719 761 Cytoplasmic (Potential).
 FT DOMAIN 201 405 Metalloprotease.
 FT DOMAIN 406 493 Disintegrin-like.
 FT DOMAIN 494 630 Cys-rich.
 FT DOMAIN 631 664 EGF-like.
 FT SITE 174 174 Cysteine switch (Potential).
 FT METAL 342 342 Zinc (catalytic) (By similarity).

ID	Q8GZ83	PRELIMINARY;	PRT; 1187 AA.
AC	Q8GZ83;		
DT	01-MAR-2003 (TReMBLrel. 23, Created)		
DT	01-MAR-2004 (TReMBLrel. 23, Last sequence update)		
DT	01-MAR-2004 (TReMBLrel. 26, Last annotation update)		
DE	Putative disease resistance protein.		
GN	Name=At5g45260/K9E15_2;		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eucots; rosids;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RA	SEQUENCE FROM N.A.		
RA	Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,		
RA	Kanajiima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,		
RA	Hayaishizaki Y., Shinozaki K.;		
RL	Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AK117156; BAC41834.1; -		
DR	GO; GO:0015020; C:membrane; IEA.		
DR	GO; GO:0005524; F:ATP binding; IEA.		
DR	GO; GO:0004888; F:transmembrane receptor activity; IEA.		
DR	GO; GO:0006915; P:apoptosis; IEA.		
DR	GO; GO:0042829; P:defense response to pathogen; IEA.		
DR	InterPro; IPRO00767; Disease_resist.		
DR	InterPro; IPRO01611; LRR.		
DR	InterPro; IPRO02182; NB-ARC.		
DR	InterPro; IPRO00157; TIR.		
DR	Pfam; PF00560; LRR; 5.		
DR	Pfam; PF00931; NB-ARC; 2.		
DR	PRINTS; PR00364; DISEASERISIT.		
SQ	SEQUENCE 1187 AA; 134249 MW; 0B1E2732280A9B31 CRC64;		
Query Match	47.68; Score 49; DB 2; Length 1187;		
Best Local Similarity	36.8%; Pred.No. 99;		
Matches	7; Conservative 9; Mismatches 3; Indels 0; Gaps		
OY	1 EDLESVLRIINWAKGSPI 19		
DD	::::: :: :: :		
	324 QNLQELSVRVINYANGNPL 342		
RESULT 8			
Q9FH84	PRELIMINARY;	PRT; 1187 AA.	
ID	Q9FH84		
AC	Q9FH84;		
DT	01-WAR-2001 (TReMBLrel. 16, Created)		
DT	01-WAR-2001 (TReMBLrel. 16, Last sequence update)		
DT	01-WAR-2004 (TReMBLrel. 26, Last annotation update)		
DE	Disease resistance protein-like.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eucots; rosids;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RA	SEQUENCE FROM N.A.		
RA	MEDLIN=20181125; PubMed=10718197;		
RX	Sato S., Nakamura Y., Kaneko T., Asamizu E., Kotani H.,		
RA	Tabata S.;		
RT	"Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequen-		
RT	features of the regions of 3,076,755 bp covered by sixty P1 and TAC		
RT	clones";		
RL	DNA Res. 7:31-63 (2000).		
DR	EMBL; AB020744; BAB10247.1; -		
DR	GO; GO:0016020; C:membrane; IEA.		
DR	GO; GO:0005524; F:ATP binding; IEA.		
DR	GO; GO:0004888; F:transmembrane receptor activity; IEA.		
DR	GO; GO:0006915; P:apoptosis; IEA.		
DR	GO; GO:0042829; P:defense response to pathogen; IEA.		
DR	InterPro; IPRO00767; Disease_resist.		
DR	InterPro; IPRO01611; LRR.		
DR	InterPro; IPRO02182; NB-ARC.		


```

RL  Nat. Biotechnol. 22:547-553(2004).
DR  EMBL; AE017222; AAS82524.1; -.
DR  InterPro; IPR000209; Pept_S8_S53.
DR  Pfam; PF00082; Peptidase_S8; 1.
DR  PRINTS; PR00723; SUBTILISIN.
DR  PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR  PROSITE; PS00138; SUBTILASE_SER; UNKNOWN_1.
KW  complete proteome; Hypothetical protein; Plasmid.
SQ  SEQUENCE 594 AA; 60515 MW; 2C95212457B8554E CRC64;

Query Match 46.6%; Score 48; DB 2; Length 594;
Best Local Similarity 52.9%; Pred. No. 66;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy  4 ESVLIRLINWAKGSPIP 20
    ||| : ||| : |||
Db  253 ESDLIALRWAGIPVP 269

RESULT 11
AAS82524 PRELIMINARY; PRT; 594 AA.
AC AAS82524;
DT 14-APR-2004 (TEMBLrel. 27, Created)
DT 14-APR-2004 (TEMBLrel. 27, Last sequence update)
DT 11-MAY-2004 (TEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN TP0194
OS Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
CX NCBI_TaxID=262724;
CX NCBI_TaxID=262724;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15064768;
RA Henne A., Brueggemann H., Raasch C., Wierzer A., Hartsch T.,
RA Liesegang H., Johann A., Lienard T., Gohl O., Martinez-Arias R.,
RA Jacobi C., Starkuviene V., Schlenczek S., Dencker S., Huber R.,
RA Klenk H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.;
RT "The genome sequence of the extreme thermophile Thermus
RT thermophilus".
RL Nat. Biotechnol. 22:547-553(2004).
DR EMBL; AE017222; AAS82524.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 594 AA; 60515 MW; 2C95212457B8554E CRC64;

Query Match 46.6%; Score 48; DB 2; Length 594;
Best Local Similarity 52.9%; Pred. No. 66;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy  4 ESVLIRLINWAKGSPIP 20
    ||| : ||| : |||
Db  253 ESDLIALRWAGIPVP 269

RESULT 12
Q8XING PRELIMINARY; PRT; 760 AA.
ID Q8XING
AC Q8XING;
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TEMBLrel. 26, Last annotation update)
DE Oxidosqualene:lanosterol cyclase.
GN Name=OSLC1;
OS Cephalosporium caerules.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; mitospore Hypocreales;
OC Cephalosporium.
CX NCBI_TaxID=150420;
CX NCBI_TaxID=150420;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21621401; PubMed=11750056;

```

RA Abe I., Naito K., Takagi Y., Noguchi H.;
RT "Molecular cloning, expression, and site-directed mutations of
RL Oxidoreductase cyclase from *Cephalosporium caerulens*.";
RL Biochim. Biophys. Acta 1522:67-73(2001).
DR ENBL: AF327881; AAL56020.1; -
DR GO: GO:001629; F:lyase activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR001330; Pfam: Pfam0008152; Trp synthase.
DR InterPro: IPR002365; Terpen synthase.
DR InterPro: IPR008930; Terp cyc toroid.
DR Pfam: PF00432; Prenyltrans; 3_
DR TIGRfam: TIGR01787; squalene_cyclase; 1.
SQ SEQUENCE 760 AA; 87082 MW; 49C01F4B50D15FE1 CRC64;

Query Match 46.6%; Score 48; DB 2; Length 760;
Best Local Similarity 42.1%; Pred. No. 87;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 EDLESVLIRLINWAKSP1 19
DB 579 EDIEKFIKRNWIRGSSL 597

RESULT 13
ID Q922R7 PRELIMINARY; PRT; 1104 AA.
AC Q922R7
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Replication factor C (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Blum A., Simon B.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR ENBL: AF030050; AAD01990.1; -
DR GO: GO:0005663; C:DNA replication factor C complex; IEA.
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0000166; F:nucleotide binding; IEA.
DR GO: GO:0006260; P:DNA replication; IEA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003959; AAA_ATPase_cent.
DR InterPro: IPR001357; BRCT.
DR InterPro: IPR000862; RFC.
DR Pfam: PF00004; AAA; 1.
DR Pfam: PF00533; BRCT; 1.
DR SMART: SM00382; AAA; 1.
DR SMART: SM00292; BRCT; 1.
DR PROSITE: PS00172; BRCT; 1.
KW ATP-binding.
FT NON TER 1104 1104
SQ SEQUENCE 1104 AA; 123246 MW; B8E1E4BCE650793B CRC64;

Query Match 46.6%; Score 48; DB 2; Length 1104;
Best Local Similarity 45.0%; Pred. No. 13e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 EDLESVLIRLINWAKSP1 20
DB 87 EKLPRVLNVLNLSGSLP 106

RESULT 14
ID P78847 PRELIMINARY; PRT; 1647 AA.
AC P78847
DT 01-MAY-1997 (TrEMBLrel. 03, Created)

DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE SPC290.03c protein.
GN Name=SPC290.03c;
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972b-;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgours J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moutle S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Welter J., Vanstreels E., Rieger M., Schafer M., Muller-Auer S.,
RA Gabel C., Fuchs M., Dusterhoft A., Fritz C., Holzer E., Moestl D.,
RA Hilbert H., Borzym K., Langer I., Beck A., Lehach H., Reinhardt R.,
RA Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Fumelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas R., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,
RA Cerutti L., Lowe T., McCombie W.K., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of *Schizosaccharomyces pombe*."
RL Nature 415:871-880(2002).
RL ENBL: AL035260; CAA22873.1; -
DR PIR: T41267; T41267.
DR PIR: T42996; T42996.
DR GeneDB Spombe; SPC290.03c; -
DR GO: GO:0005635; C:nuclear membrane; IEA.
DR GO: GO:0000917; P:barrier septum formation; IMP.
DR GO: GO:0000917; P:barrier septum formation; IMP.
SQ SEQUENCE 1647 AA; 186410 MW; 579B36CB585ED618 CRC64;

Query Match 46.6%; Score 48; DB 2; Length 1647;
Best Local Similarity 40.0%; Pred. No. 2.1e+02;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 EDLESVLIRLINWAKSP1 20
DB 477 DDIESDMVGLTWSGSLP 496

RESULT 15
ID Q9SF12 PRELIMINARY; PRT; 554 AA.
AC Q9SF12
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein F26K24.16.
GN Name=F26K24.16;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; rosids;
OC Eucotyledons II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.,

RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC016795; AAF23203.1; -;
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0005468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_Kinase.
DR InterPro; IPR006567; PUG.
DR InterPro; IPR010513; Ribonuc_2-5A.
DR Pfam; PF00069; Kinase; 1.
DR Pfam; PF06479; Ribonuc_2-5A; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00580; PUG; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
KW Hypothetical protein.
SQ SEQUENCE 554 AA, 63259 MW, 3543FFF1A6651683 CRC64;

Query Match 46.1%; Score 47.5; DB 2; Length 554;
Best Local Similarity 47.6%; Pred. NC. 73;
Matches 10; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 EDLESVLRLNW-AKGSPIP 20
Db 224 DDLEKVMKRIKFWKEGRPLP 244

Search completed: November 10, 2004, 14:50:22
Job time : 28.6415 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 9.15094 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-57

Perfect score: 137

Sequence: 1 RPSVFCGAVWTLNRAIGHFVRGSR 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	49	35.8	376	1	US-08-002-024B-8
2	48	35.0	919	4	US-08-583-681A-7187
3	47	34.3	995	4	US-09-252-991A-22297
4	46	33.6	464	4	US-09-252-991A-27800
5	45	32.8	252	4	US-09-252-991A-26127
6	45	32.8	452	4	US-09-489-039A-9355
7	45	32.8	607	3	US-08-969-683A-59
8	45	32.8	1693	3	US-08-478-507-7
9	45	32.8	1693	3	US-08-128-275A-7
10	45	32.8	1693	3	US-09-553-427-7
11	44.5	32.5	107	4	US-09-673-395A-392
12	44.5	32.5	175	4	US-09-270-767-41770
13	44	32.1	63	2	US-08-637-759B-442
14	44	32.1	63	3	US-08-871-355A-442
15	44	32.1	63	3	US-09-201-945-442
16	44	32.1	235	4	US-09-489-039A-7572
17	44	32.1	237	4	US-09-489-039A-9365
18	44	32.1	400	4	US-09-134-000C-4389
19	44	32.1	1708	4	US-09-462-606-2
20	43.5	31.8	378	2	US-08-846-762-87
21	43.5	31.8	405	4	US-09-489-039A-9838
22	43	31.4	371	3	US-09-193-069-2
23	43	31.4	455	4	US-09-248-796A-16498
24	43	31.4	528	4	US-10-140-372-7
25	43	31.4	730	2	US-08-696-944-2
26	42.5	31.0	206	4	US-09-270-767-57797
27	42.5	31.0	274	4	US-09-270-767-42500, A

28	42.5	31.0	470	4	US-09-248-796A-15129	Sequence 15129, A
29	42.5	31.0	799	4	US-09-489-039A-8219	Sequence 8219, Ap
30	42	30.7	59	4	US-09-513-999C-6346	Sequence 6346, Ap
31	42	30.7	60	4	US-09-540-236-3288	Sequence 3288, Ap
32	42	30.7	181	4	US-09-270-767-40764	Sequence 40764, A
33	42	30.7	181	4	US-09-270-767-55980	Sequence 55980, A
34	42	30.7	162	4	US-09-252-991A-26062	Sequence 26062, A
35	42	30.7	184	4	US-09-252-991A-28524	Sequence 28524, A
36	42	30.7	460	4	US-09-489-039A-13680	Sequence 12680, A
37	42	30.7	477	1	US-07-735-065-2	Sequence 2, Appli
38	42	30.7	477	1	US-08-469-202-12	Sequence 12, Appli
39	42	30.7	477	2	US-08-484-434C-12	Sequence 12, Appli
40	42	30.7	477	4	US-09-384-361-12	Sequence 12, Appli
41	42	30.7	580	3	US-08-818-112-75	Sequence 75, Appli
42	42	30.7	580	3	US-08-818-111-76	Sequence 76, Appli
43	42	30.7	580	3	US-09-056-556-75	Sequence 75, Appli
44	42	30.7	580	4	US-09-072-596-76	Sequence 76, Appli
45	42	30.7	580	4	US-09-072-967-75	Sequence 75, Appli

ALIGNMENTS

RESULT 1
US-08-002-024B-8
; Sequence 8, Application US/08002024B
; Patent No. 5798103
; GENERAL INFORMATION:
; APPLICANT: MOOI, Frederik R
; TITLE OF INVENTION: WHOOPING COUGH VACCINE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: YOUNG & THOMPSON
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: VA
; COUNTRY: US
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/002,024B
; FILING DATE: 08-JAN-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 92200038.5
; FILING DATE: 08-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Robert J
; REGISTRATION NUMBER: 17,355
; REFERENCE/DOCKET NUMBER: BO-37424
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703/521-2297
; TELEFAX: 703/685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-002-024B-8

Query Match 35.8%; Score 49; DB 1; Length 376;
Best Local Similarity 34.2%; Pred. No. 9.1;
Matches 13; Conservative 2; Mismatches 5; Indels 18; Gaps 2;

QY 1 RPSVFCG---AVMT-----LNRAIGHF 20
DB 15 RVISFCGALAWAGLVQPMNAVDPVDCGALGLHF 52

```
RESULT 2
US-09-543-681A-7187
; Sequence 7187, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETTON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7187
; LENGTH: 919
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7187

Query Match 35.0%; Score 48; DB 4; Length 919;
Best Local Similarity 41.7%; Pred. No. 37;
Matches 10; Conservative 4; Mismatches 6; Indels 4; Gaps 1;

QY 1 RPVSFCGAVWTLNRAIGHRVGRS 24
|||:||||:|:|:|:
Db 494 RPVAFCGG---TSVAGYIVRG 513

RESULT 3
US-09-252-991A-22297
; Sequence 22297, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22297
; LENGTH: 995
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22297

Query Match 34.3%; Score 47; DB 4; Length 995;
Best Local Similarity 52.6%; Pred. No. 58;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 4 SFCGAVWTLNRAIGHRVGR 22
|||:||||:|:|:|:
Db 160 SFGALWLFERAVVRAGR 178

RESULT 4
US-09-252-991A-27800
; Sequence 27800, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
```

```
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27800
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27800

Query Match 33.6%; Score 46; DB 4; Length 464;
Best Local Similarity 52.6%; Pred. No. 35;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 PVSFCGAVWTLNRAIGHRF 20
|||:||||:|:|:|:
Db 115 PVLATGATYLNALGTHF 133

RESULT 5
US-09-252-991A-26127
; Sequence 26127, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26127
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26127

Query Match 32.8%; Score 45; DB 4; Length 252;
Best Local Similarity 66.7%; Pred. No. 25;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 14 RAIGHRVGRS 25
|||:||||:|:|:|:
Db 80 RATGRHWARGTR 91

RESULT 6
US-09-489-039A-9355
; Sequence 9355, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Bretton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9355
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9355

Query Match 32.8%; Score 45; DB 4; Length 452;
Best Local Similarity 56.2%; Pred. No. 49;
```

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 VSFCGAVWTLNRAICR 18
DB 145 VDFLDVWVWLNALDR 160

RESULT 7
US-08-969-683A-59
; Sequence 59, Application US/08969683A
; Patent No. 6136576
; GENERAL INFORMATION:
; APPLICANT: GENENCOR INTERNATIONAL, INC.
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
; PRODUCTION OF 1,3 PROPANEDIOL
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International, Inc.
; STREET: 4 Cambridge Place
; STREET: 1870 South Winton Road
; CITY: Rochester
; STATE: NY
; COUNTRY: U.S.A
; ZIP: 14618
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/969,683A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/20873
; FILING DATE: 13-NOV-1997
; APPLICATION NUMBER: 60/030,601
; FILING DATE: 13-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Glaister, Debra
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: GC 369-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-864-7620
; TELEFAX: 650-845-6504
; TELEX:
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 607 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6136576e
US-08-969-683A-59

Query Match 32.8%; Score 45; DB 3; Length 607;
Best Local Similarity 56.2%; Pred. No. 69;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 VSFCGAVWTLNRAICR 18
DB 141 VDFLDVWVWLNALDR 156

RESULT 8
US-08-478-507-7
; Sequence 7, Application US/08478507
; Patent No. 6120988
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; APPLICANT: Yarbough, Patrice O
; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z

; APPLICANT: Tam, Albert
; APPLICANT: Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; TITLE OF INVENTION: No. 6120988-A/No. 6120988-B Hepatitis Viral Agent
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,507
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0183.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1693 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-478-507-7

Query Match 32.8%; Score 45; DB 3; Length 1693;
Best Local Similarity 38.1%; Pred. No. 2,2e+02;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 4 SFCGAVWTLNRAIGHFVRGS 24
DB 647 AFCSALYRFNREARHSLIGN 667

RESULT 9
US-09-128-275A-7
; Sequence 7, Application US/09128275A
; Patent No. 6229005
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; APPLICANT: Yarbough, Patrice O
; APPLICANT: Bradley, Daniel W

```

; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; TITLE OF INVENTION: NO. 6229005-A/No. 6229005-B Hepatitis Viral Agent
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/128,275A
; FILING DATE: 03-AUG-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Pettithory, Joane R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 4600-0183.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1693 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-128-275A-7

```

```

Query Match 32.8%; Score 45; DB 3; Length 1693;
Best Local Similarity 38.1%; Pred. No. 2.2e+02;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 4 SFCGAVWTLNRAIGHFVRGS 24
Db 647 AFCSALYRFNREARHSLIGN 667

```

```

RESULT 10
US-09-553-427-7
; Sequence 7, Application US/09553427
; Patent No. 6379891
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; APPLICANT: Yarbrough, Patrice O

```

```

; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; TITLE OF INVENTION: NO. 6379891-A/No. 6379891-B Hepatitis Viral Agent
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/553,427
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/478,507
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0183.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1693 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-553-427-7

```

```

Query Match 32.8%; Score 45; DB 3; Length 1693;
Best Local Similarity 38.1%; Pred. No. 2.2e+02;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 4 SFCGAVWTLNRAIGHFVRGS 24
Db 647 AFCSALYRFNREARHSLIGN 667

```

```

RESULT 11
US-09-673-395A-392
; Sequence 392, Application US/09673395A
; Patent No. 6620923

```



```
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
; FILE REFERENCE: ALBRE-12
; CURRENT APPLICATION NUMBER: US/09/673,395A
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 637
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 392
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-673-395A-392

Query Match      32.5%; Score 44.5; DB 4; Length 107;
Best Local Similarity 47.4%; Pred. No. 11;
Matches 9; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY      2 PVS-FCGAVWTLNRAIGHF 19
      |:|:|:|:|:|:|:|:|:|:|
DB      28 PITQFCFIWTPSLQGHF 46

RESULT 12
US-09-270-767-41770
; Sequence 41770, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41770
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-41770

Query Match      32.5%; Score 44.5; DB 4; Length 175;
Best Local Similarity 42.9%; Pred. No. 20;
Matches 9; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY      1 RPVSFCGAVWTLNRAIGHFV 21
      |:|:|:|:|:|:|:|:|:|:|
DB      58 RPPWMC---WRPSPGKGRHF 75

RESULT 13
US-08-637-759B-442
; Sequence 442, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101 CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 442:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 63 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
US-08-637-759B-442

Query Match      32.1%; Score 44; DB 2; Length 63;
Best Local Similarity 58.3%; Pred. No. 7.5;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      9 VWTLNRAIGHF 20
      |||||:|:|:|:|:|:|:|
DB      42 VWTLNNSCTRH 53

RESULT 14
US-08-871-355A-442
; Sequence 442, Application US/08871355A
; Patent No. 6015669
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,355A
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101 CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
```

INFORMATION FOR SEQ ID NO: 442:

SEQUENCE CHARACTERISTICS:
LENGTH: 63 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-871-355A-442

Query Match 32.1%; Score 44; DB 3; Length 63;
Best Local Similarity 58.3%; Pred. No. 7.5;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 9 VMTLNRAIGHF 20
Db 42 VMTLNNSCTHY 53

RESULT 15

US-09-201-945-442
Sequence 442, Application US/09201945
Patent No. 6342215

GENERAL INFORMATION:

APPLICANT: David William Holden

TITLE OF INVENTION: Identification of Genes

NUMBER OF SEQUENCES: 501

CORRESPONDENCE ADDRESS:

ADDRESSEE: Patrea L. Pabst

STREET: 2800 One Atlantic Center

CITY: Atlanta

STATE: Georgia

COUNTRY: USA

ZIP: 30309-3450

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/201,945

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/637,759

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Pabst, Patrea L.

REGISTRATION NUMBER: 31,284

REFERENCE/DOCKET NUMBER: RPMS 101

TELECOMMUNICATION INFORMATION:

TELEPHONE: (404) 873-8794

TELEFAX: (404) 873-8795

INFORMATION FOR SEQ ID NO: 442:

SEQUENCE CHARACTERISTICS:

LENGTH: 63 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

US-09-201-945-442

Query Match 32.1%; Score 44; DB 3; Length 63;
Best Local Similarity 58.3%; Pred. No. 7.5;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 9 VMTLNRAIGHF 20
Db 42 VMTLNNSCTHY 53

Search completed: November 10, 2004, 14:55:41
Job time : 10.2009 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 28.9151 Seconds
(without alignment)
305.399 Million cell updates/sec

Title: US-10-092-750-57

Perfect score: 137
Sequence: 1 RPVSFCGAVTTLNRAIGRHFVRGSR 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 35322586 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	137	100.0	25	14	US-10-092-750-57
2	56.5	41.2	231	17	US-10-425-115-234613
3	53	38.7	168	16	US-10-437-963-163106
4	51	37.2	3042	16	US-10-437-963-123736
5	50.5	36.9	932	14	US-10-444-575-4
6	50	36.5	239	17	US-10-425-115-270985
7	49	35.8	1707	16	US-10-239-090A-12
8	49	35.8	1707	16	US-10-239-090A-54
9	49	35.8	1707	16	US-10-239-090A-56
10	47.5	34.7	91	17	US-10-425-115-256918
11	47.5	34.7	214	17	US-10-425-115-302062
12	47.5	34.7	536	17	US-10-425-115-187602
13	47	34.3	107	15	US-10-424-599-246506

14	47	34.3	235	15	US-10-424-599-246502	Sequence 246502,
15	47	34.3	247	15	US-10-425-114-40659	Sequence 40659, A
16	47	34.3	312	15	US-10-425-114-49039	Sequence 49039, A
17	47	34.3	393	15	US-10-425-114-43621	Sequence 43621, A
18	47	34.3	493	15	US-10-424-599-246507	Sequence 246507,
19	47	34.3	727	15	US-10-425-114-52764	Sequence 52764, A
20	47	34.3	727	15	US-10-425-114-55408	Sequence 55408, A
21	46	33.6	106	11	US-10-424-599-235252	Sequence 235252,
22	46	33.6	104	11	US-09-864-408A-4608	Sequence 4608, Ap
23	46	33.6	117	17	US-10-425-115-23191	Sequence 25191,
24	46	33.6	348	15	US-10-424-599-159122	Sequence 159122,
25	46	33.6	394	15	US-10-425-114-37391	Sequence 37391, A
26	46	33.6	394	14	US-10-369-493-12891	Sequence 12891, A
27	46	33.6	469	14	US-10-369-493-22836	Sequence 22836, A
28	46	33.6	536	16	US-10-437-963-162720	Sequence 162720,
29	46	33.6	551	14	US-10-223-277-3	Sequence 3, Appli
30	45.5	32.2	200	17	US-10-425-115-302416	Sequence 302416,
31	45	32.8	116	16	US-10-767-701-50399	Sequence 50399, A
32	45	32.8	161	14	US-10-029-386-33432	Sequence 33432, A
33	45	32.8	212	15	US-10-424-599-272781	Sequence 272781,
34	45	32.8	269	15	US-10-424-599-234166	Sequence 234166,
35	45	32.8	343	14	US-10-369-493-1046	Sequence 1046, Ap
36	45	32.8	438	14	US-10-156-761-13970	Sequence 13970, A
37	45	32.8	478	15	US-10-320-797-3206	Sequence 3206, Ap
38	45	32.8	541	9	US-09-815-242-10746	Sequence 10746, A
39	45	32.8	607	10	US-09-308-207-59	Sequence 59, Appli
40	45	32.8	1693	10	US-09-851-410-7	Sequence 7, Appli
41	44.5	32.5	69	16	US-10-437-963-160071	Sequence 160071,
42	44.5	32.5	86	17	US-10-425-115-196350	Sequence 196350,
43	44.5	32.5	97	17	US-10-425-115-264220	Sequence 264220,
44	44.5	32.5	160	15	US-10-424-599-238559	Sequence 238559,
45	44.5	32.5	463	15	US-10-389-586-1198	Sequence 1198, Ap

ALIGNMENTS

RESULT 1
US-10-092-750-57
; Sequence 57, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-57

Query Match 100.0%; Score 137; DB 14; Length 25;
Best Local Similarity 100.0%; Pred. No. 7.1e-14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPVSFCGAVTTLNRAIGRHFVRGSR 25

Db 1 RPVSFCGAVTTLNRAIGRHFVRGSR 25

RESULT 2
US-10-425-115-234613
; Sequence 234613, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:

```

Query Match      36.9%; Score 50.5; DB 14; Length 932;
Best local Similarity 28.8%; Pred. No. 71;
Matches 11; Conservative 6; Mismatches 8; Indels 13; Gaps 1;
Qy      1 RPVSFCGAVWTLNRAIGRHF-----VRGSR 25
      : |||: |||: |||: |||:

```



```
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(91)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_165898C.1.pep
US-10-425-115-256918

Query Match      34.7%; Score 47.5; DB 17; Length 91;
Best Local Similarity 57.9%; Pred. No. 18;
Matches 11; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

QY 1 RVPVFCGAVW-TLNRAIGR 18
Db 17 RLVPVCGAVWLVLGGAVGR 35

RESULT 11
US-10-425-115-302062
; Sequence 302062, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 302062
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_36556C.1.pep
US-10-425-115-302062

Query Match      34.7%; Score 47.5; DB 17; Length 214;
Best Local Similarity 44.0%; Pred. No. 44;
Matches 11; Conservative 4; Mismatches 7; Indels 3; Gaps 1;

QY 1 RVPVFCGAVW-TLNRAIGRHFVR 22
Db 147 RPAHECSCGVSFTSCALGGHWR 171

RESULT 12
US-10-425-115-187602
; Sequence 187602, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 187602
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(536)

; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_102679C.1.pep
US-10-425-115-187602

Query Match      34.7%; Score 47.5; DB 17; Length 536;
Best Local Similarity 37.0%; Pred. No. 1.2e+02;
Matches 10; Conservative 3; Mismatches 9; Indels 5; Gaps 1;

QY 2 PVSCGAVWTLNRAIGRHF-----VRG 23
Db 240 PIMFCSLKTQNRXSRHFFHLHCIRG 266

RESULT 13
US-10-424-599-246506
; Sequence 246506, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 246506
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_64626C.1.pep
US-10-424-599-246506

Query Match      34.3%; Score 47; DB 15; Length 107;
Best Local Similarity 57.1%; Pred. No. 26;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 GAVWTLNRAIGRHF 20
Db 26 GEVWNGRSIGRHW 39

RESULT 14
US-10-424-599-246502
; Sequence 246502, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 246502
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(235)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_64622C.1.pep
US-10-424-599-246502
```

```

Query Match      34.3%; Score 47; DB 15; Length 235;
Best Local Similarity 57.1%; Pred. No. 58;
Matches      8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      7 GAVWTLNRAIGRHF 20
Db      154 GEVWNGRSIGRHW 167

RESULT 15
US-10-425-114-40659
; Sequence 40659, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 40659
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700846835_FLI.pep
US-10-425-114-40659

Query Match      34.3%; Score 47; DB 15; Length 247;
Best Local Similarity 57.1%; Pred. No. 61;
Matches      8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      7 GAVWTLNRAIGRHF 20
Db      166 GEVWNGRSIGRHW 179

Search completed: November 11, 2004, 02:43:13
Job time : 29.9651 secs

```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 6.13208 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-57

Perfect score: 137

Sequence: 1 RPVSFCGAVMTLNRAIGRHFRVGRSR 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	51	37.2	376	2	S42746	fimbrial adhesin f
2	50.5	36.9	932	2	JC5953	inter-alpha-inhibi
3	50	36.5	496	2	T15691	hypothetical prote
4	49	35.8	376	2	S42747	fimbrial adhesin f
5	49	35.8	376	2	S36247	fimbrial adhesin f
6	48	35.0	195	2	T32807	hypothetical prote
7	48	35.0	366	2	T51401	hypothetical prote
8	48	35.0	407	2	D87534	hypothetical prote
9	48	35.0	557	2	T41588	probable amino-aci
10	47	34.3	318	2	A86750	hypothetical prote
11	46	33.6	193	2	AH1855	hypothetical prote
12	46	33.6	237	2	S14722	hypothetical prote
13	46	33.6	273	2	S69193	probable finger pr
14	46	33.6	455	2	F83460	probable amino aci
15	46	33.6	469	2	D72722	probable MESA prot
16	46	33.6	472	2	AH0109	probable histidine
17	45.5	33.2	362	2	C82070	conserved hypotnet
18	45	32.8	256	2	E70868	probable enoyl-coA
19	45	32.8	315	2	S44019	aerobactin biosynt
20	45	32.8	315	2	T44480	conserved hypotnet
21	45	32.8	340	2	E81331	uroporphyrinogen d
22	45	32.8	343	2	A64443	glyceraldehyde-3-p
23	45	32.8	447	2	A81919	probable sodium-tr
24	45	32.8	529	2	JE0181	oligo-1,6-glucosid
25	45	32.8	668	2	AH1565	hypothetical prote
26	45	32.8	1693	1	MNWWEH	genome polyprotein
27	44.5	32.5	351	2	A88103	protein WI0G11.7 l
28	44.5	32.5	463	2	E95277	probable OtsA treh
29	44.5	32.5	1197	2	D82696	hypothetical prote

30	44	32.1	147	2	S70833	methylated-DNA-pro
31	44	32.1	229	1	TRDFS	trypsin (EC 3.4.21
32	44	32.1	330	2	T25939	hypothetical prote
33	44	32.1	564	2	H75403	glycosyl hydrolase
34	43.5	31.8	351	2	H70682	probable cysA - My
35	43.5	31.8	353	2	A25580	photosystem II pro
36	43.5	31.8	559	1	A35029	t-plasminogen acti
37	43.5	31.8	628	2	A70376	conserved hypotnet
38	43	31.4	157	2	H70317	glycerol-3-phospha
39	43	31.4	207	1	G45340	hypothetical prote
40	43	31.4	257	2	T31609	hypothetical prote
41	43	31.4	281	2	S19159	finger protein EPP
42	43	31.4	284	2	A84682	probable zinc-fing
43	43	31.4	326	2	B90841	probable ABC trans
44	43	31.4	326	2	A85699	probable ABC trans
45	43	31.4	339	2	E71918	uroporphyrinogen d

ALIGNMENTS

RESULT 1

S42746
fimbrial adhesin fimD - Bordetella parapertussis
C:Species: Bordetella parapertussis
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S42746
R:Willems, R.J.L.
submitted to the EMBL Data Library, October 1993
A:Reference number: S42746
A:Accession: S42746
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-376 <WIL>
A:Cross-references: UNIPROT:Q44888; EMBL:X75812; NID:G416070; PIDN:CAA53447.1; PID:G416

Query Match 37.2%; Score 51; DB 2; Length 376;
Best Local Similarity 36.8%; Pred. No. 4.4;
Matches 14; Conservative 1; Mismatches 5; Indels 18; Gaps 2;

Qy 1 RPVSFCG---AVWT-----LNRAIGRHF 20
|:|||||
Db 15 RVISFCGALAVWAGLAVQFAMAVDPVDCGRLGHLHF 52

RESULT 2

JC5953
inter-alpha-inhibitor H4P heavy chain - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004
C:Accession: JC5953
R:Scoury, E.; Olivier, E.; Daveau, M.; Hiron, M.; Claeysens, S.; Risler, J.L.; Salier, Biochem. Biophys. Res. Commun. 243, 522-530, 1998
A:Title: The H4P heavy chain of inter-alpha-inhibitor family largely differs in the str
A:Reference number: JC5953
A:Accession: JC5953
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-932 <SQU>
A:Cross-references: UNIPROT:Q35802; GB:Y11283; NID:G2292987; PIDN:CAA72155.1; PID:G22929
C:Superfamily: inter-alpha-trypsin inhibitor complex component II

Query Match 36.9%; Score 50.5; DB 2; Length 932;
Best Local Similarity 28.9%; Pred. No. 13;
Matches 11; Conservative 6; Mismatches 8; Indels 13; Gaps 1;

Qy 1 RPVSFCGAVMTLNRAIGRHF-----VRGSR 25
|:|||||
Db 606 KPTVEDGVWVSIISAVQRHFKPTTSGKLTSLRIGNR 643

RESULT 3

T15691

A/Molecule type: DNA
A/Residues: 1-366 <SAT>
A/Cross-references: UNIPROT:Q9LFV4; EMBL:AL391144
A/Experimental source: cultivar Columbia; BAC clone F14F8
C/Genetics:
A/Map position: 5
A/Introns: 80/1; 124/3; 207/3; 294/1
A/Notes: F14F8_100
C/Superfamily: myrosinase-associated protein MyAP

Query Match 35.0%; Score 48; DB 2; Length 366;
Best Local Similarity 40.0%; Pred. No. 12;
Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 RPVSFCGAVVTLNRAIGSHFVRGSR 25
Db 103 RGVNYASAAAGILDTGRHYVRGAR 127

RESULT 8
D87534
hypothetical protein CC2301 [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: D87534
R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: D87534
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-407 <STO>
A/Cross-references: UNIPROT:Q9A5Z6; GB:AEO05673; NID:gl3423820; PIDN:AAK24272.1; GSPDB:J.
C/Genetics:
A/Gene: CC2301

Query Match 35.0%; Score 48; DB 2; Length 407;
Best Local Similarity 52.9%; Pred. No. 14;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 7 GAVVTLNRAIGRHVRVG 23
Db 229 GVWLGTRVICKPFVRG 245

RESULT 9
T41588
probable amino-acid permease - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T41588
R/Murphy, L.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1998
A/Reference number: Z21968
A/Accession: T41588
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-557 <MUR>
A/Cross-references: UNIPROT:O74537; EMBL:AL031543; PIDN:CAA20834.1; GSPDB:GN000669; SPDB:
A/Experimental source: strain 972h-; cosmid c74
C/Genetics:
A/Gene: SPDB:SPCC74.04
A/Map position: 3
C/Superfamily: choline transport protein

Query Match 35.0%; Score 48; DB 2; Length 557;
Best Local Similarity 54.5%; Pred. No. 19;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 VSFCGAVVTLN 13

A;Accession: S14722

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-237 <FAN>

A;Cross-references: GB:X59504; NID:G561; PIDN:CAA42092.1; PID:G562

C;Superfamily: prolactin

Query Match 33.6%; Score 46; DB 2; Length 237;

Best Local Similarity 52.4%; Pred. No. 16;

Matches 11; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

QY 4 SFCGAVVTLNRAIGRHFVRGS 24

||||| |||

Db 6 SFCGHWQTYNP-----VRGS 20

RESULT 13

S69193

probable finger protein Pszf1 - garden pea

C;Species: Pisum sativum (garden pea)

C;Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

C;Accession: S69193

R;Michael, A.J.; Hofer, J.M.I.; Ellis, T.H.N.

Plant Mol. Biol. 30, 1051-1058, 1996

A;Title: Isolation by PCR of a cDNA clone from pea petals with similarity to petunia and

A;Reference number: S69193; MUID:96270382; PMID:8639742

A;Accession: S69193

A;Molecule type: mRNA

A;Residues: 1-273 <MIC>

A;Cross-references: UNIPROT:Q41070; EMBL:X87374; NID:G854399; PIDN:CAA60828.1; PID:G8610

C;Superfamily: Arabidopsis thaliana hypothetical protein F1ZE4.290

C;Keywords: DNA binding; zinc finger

F;115-135/Region: zinc finger CCH motif

F;202-222/Region: zinc finger CCH motif

Query Match 33.6%; Score 46; DB 2; Length 273;

Best Local Similarity 47.4%; Pred. No. 19;

Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 SFCGAVVTLNRAIGRHFVR 22

||||| :|||

Db 203 SICGAEFTSGQALGGHVR 221

RESULT 14

F83460

probable amino acid permease PA1485 [imported] - Pseudomonas aeruginosa (strain PA01)

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C;Accession: F83460

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: F83460

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-455 <STO>

A;Cross-references: UNIPROT:Q913M7; GB:AE004577; GB:AE004091; NID:G9947430; PIDN:AAG0487

C;Experimental source: strain PA01

C;Genetics:

A;Gene: PA1485

Query Match 33.6%; Score 46; DB 2; Length 455;

Best Local Similarity 52.6%; Pred. No. 31;

Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 PVSFCGAVVTLNRAIGRHF 20

||||| |||

Db 109 PVLATGAATYLNALGTHF 127

RESULT 15

D72722

probable MRSa protein APE0317 - Aeropyrum pernix (strain K1)

C;Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C;Accession: D72722

R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-DO, K.; Takai

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; I

DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy

A;Reference number: A72450; MUID:99310339; PMID:10382966

A;Accession: D72722

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-469 <KAW>

A;Cross-references: UNIPROT:Q9YFC4; DBJ:AF000059; NID:G5103911; PIDN:BAA79272.1; PID:G

A;Experimental source: strain K1

C;Genetics:

A;Gene: APE0317

C;Superfamily: phosphomannomutase

Query Match 33.6%; Score 46; DB 2; Length 469;

Best Local Similarity 64.7%; Pred. No. 32;

Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 9 VVTLNRAIGRHFVRGSR 25

||||| :|||

Db 30 VLRLARAIGAYFGRGSR 46

Search completed: November 10, 2004, 14:52:25

Job time : 7.13208 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 33.3019 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-57
Perfect score: 137
Sequence: 1 RPVSFCGAVTINRAIGRHFVRGSR 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt.02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	38.7	84	2	Q8KB37 chlorobium
2	52	38.0	791	2	Q8Z223 pyrobaculum
3	52	38.0	1693	2	Q9W115 hepatitis e
4	52	38.0	2503	2	Q8B912 porcine rep
5	51	37.2	365	2	Q7W695 bordetella
6	51	37.2	376	2	Q44888 bordetella
7	51	37.2	1693	2	Q69418 hepatitis e
8	50.5	36.9	932	2	Q35802 rattus norv
9	50	36.5	199	2	Q7Q1J4 anopheles g
10	50	36.5	496	2	Q18288 caenorhabdi
11	50	36.5	1706	2	Q7TGF2 hepatitis e
12	49.5	36.1	573	2	Q9V516 drosophila
13	49	35.8	341	2	P93714 petunia hyb
14	49	35.8	365	2	Q7W164 bordetella
15	49	35.8	376	1	FAAE_BORPE
16	49	35.8	376	2	Q44887 bordetella
17	49	35.8	389	2	Q6GXX0 swine hepat
18	49	35.8	389	2	Q6GXX1 swine hepat
19	49	35.8	389	2	Q6GXX2 swine hepat
20	49	35.8	389	2	Q6GXX3 swine hepat
21	49	35.8	1684	2	Q8JJM2 hepatitis e
22	49	35.8	1707	2	Q6PMR4 swine hepat
23	49	35.8	1707	2	Q7TA93 hepatitis e
24	49	35.8	1707	2	Q806D8 hepatitis e
25	49	35.8	1707	2	Q806E1 swine hepat
26	49	35.8	1707	2	Q801R6 hepatitis e
27	49	35.8	1707	2	Q8JJN3 hepatitis e
28	49	35.8	1707	2	Q8JJN7 hepatitis e
29	49	35.8	1707	2	Q9IVZ9 hepatitis e
30	49	35.8	1707	2	AAT02424 swine hep
31	48	35.0	90	2	Q9WD34 scenedesmus

32 48 35.0 165 2 Q8GIM9 uncultured
33 48 35.0 366 2 Q9LFV4 arabidopsis
34 48 35.0 407 2 Q9ASZ6 caulobacter
35 48 35.0 492 2 Q8FUF9 bruceella su
36 48 35.0 512 2 Q9JH69 turkey astr
37 48 35.0 523 2 Q8WOH1 oryza sativ
38 48 35.0 557 1 YCQ4 SCHPO
39 48 35.0 1199 2 Q7PS03 schizosacch
40 47 34.3 152 2 Q84EA5 anopheles g
41 47 34.3 223 2 Q88Y79 uncultured
42 47 34.3 262 2 Q89R29 lactobacill
43 47 34.3 318 2 Q9CGU1 bradyrhizob
44 47 34.3 424 2 Q7UWK8 lactococcus
45 47 34.3 452 2 Q82SE3 rhodospirill
Q82ss3 nitrosomona

ALIGNMENTS

RESULT 1

Q8KB37 PRELIMINARY; PRT; 84 AA.
ID Q8KB37
AC Q8KB37
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein CT1960.
GN OrderedLocusNames=CT1960;
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobaculum.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901; DOI=10.1073/pnas.132181499;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.P., Wu M.,
RA Dodson R.J., DeBoy R.T., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.F., Yang F.,
RA Holt I.E., Umayam L.A., Mason T.M., Brenner M., Shea T.P.,
RA Parksey D.S., Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.E.,
RA Radune D., Vamathevan J.J., Khouri H.M., White O., Gruber T.M.,
RA Ketchum K.A., Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RT photosynthetic, anaerobic, green-sulfur bacterium."
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514 (2002).
EMBL; AE012946; AAM73179.1; -
TIGR; CT1960; -
DR InterPro; IPR001395; Aldo/ket red.
DR Pfam; PF00248; Aldo ket red; 1.
DR ProDom; PD000288; Aldo/Ket red; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 84 AA; 9282 MW; 7246A595F4A9CE56 CRC64;

Query Match 38.7%; Score 53; DB 2; Length 84;
Best Local Similarity 57.9%; Pred. No. 3.5;
Matches 11; Conservative 2; Mismatches 4; Indels 2; Gaps 1;
Qy 1 RPVSF--CGAVTINRAIG 17
Db 33 KPLVFTKCGLVWDENRAIG 51

RESULT 2

Q8Z223 PRELIMINARY; PRT; 791 AA.
ID Q8Z223
AC Q8Z223
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Hypothetical protein PA30478.
GN OrderedLocusNames=PA30478;
OS Pyrobaculum aerophilum.

OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
 OC Thermoproteaceae; Pyrobaculum.
 OX NCBI_TaxID=13773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
 RX MEDLINE=21664397; PubMed=11792869;
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
 RA Miller J.H.;
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
 RT aerophilum"; Acad. Sci. U.S.A. 99:984-989(2002).
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
 DR EMBL; AE009769; AAL62818.1; --
 DR InterPro; IPR006025; Pept M Zn_BS.
 DR PROSITE; PS00142; ZINC PROTEASE; UNKNOWN_1.
 KW Complete proteome; Hypochemical protein.
 SQ SEQUENCE 791 AA; 86781 MW; 1D29FB59F5394DCF CRC64;
 Query Match 38.0%; Score 52; DB 2; Length 791;
 Best Local Similarity 57.9%; Pred. No. 44;
 Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
 QY 7 GAVWTLNRAIGHFVRGRS 25
 DB 550 GAVWTSNAVVAVNGSR 568
 RESULT 3
 QSWLL5 PRELIMINARY; PRT; 1693 AA.
 AC QSWLL5;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Nonstructural polypeptide.
 OS Hepatitis E virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;
 OC Hepatitis E-like viruses.
 OX NCBI_TaxID=12461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN="TK15/92;
 RX MEDLINE=99049628; PubMed=9833882;
 RA Gouvea V., Snellings N., Popek M.J., Longer C.F., Innis B.L.;
 RT "Hepatitis E virus: complete genome sequence and phylogenetic analysis
 RT of a Nepali isolate."; J.
 RL Virus Res. 57:21-26(1998).
 DR EMBL; AF051930; AAC97186.1; --
 DR GO; GO:0008174; F:RNA methyltransferase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003724; F:RNA helicase activity; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0006396; P:RNA processing; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0016032; P:viral life cycle; IEA.
 DR GO; GO:0019082; P:viral protein processing; IEA.
 DR InterPro; IPR002589; Alpp.
 DR InterPro; IPR008748; Peptidase C41.
 DR InterPro; IPR001788; RNA_dep_RNapi2.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR000606; Viral helicase1.
 DR InterPro; IPR002588; V methyltrans.
 DR Pfam; PF01661; Alpp; 1.
 DR Pfam; PF05417; Peptidase C41; 1.
 DR Pfam; PF00978; RNA_dep_RNapi2; 1.
 DR Pfam; PF01443; Viral_helicase1; 1.
 DR Pfam; PF01660; Vmethyltransf; 1.
 DR SMART; SM00506; Alpp; 1.
 KW Polyprotein.
 SQ SEQUENCE 1693 AA; 185349 MW; A895ACD1DAE2FBBD CRC64;

Query Match 38.0%; Score 52; DB 2; Length 1693;
 Best Local Similarity 42.9%; Pred. No. 92;
 Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
 QY 4 SFGCAVWTLNRAIGHFVRGS 24
 DB 647 AFCSLYRFNRAAQRHSLTGN 667
 RESULT 4
 Q8B912 PRELIMINARY; PRT; 2503 AA.
 ID Q8B912
 AC Q8B912; 2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Proteinase.
 GN Names:ORF1a;
 OS Porcine reproductive and respiratory syndrome virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Arteriviridae; Arterivirus.
 OX NCBI_TaxID=28344;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gao Z.Q., Gao X., Yang H.C.;
 RT "Genomic characterization of two Chinese isolates of Porcine
 RT respiratory and reproductive syndrome virus";
 RL Arch. Virol. 149:1341-1351(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Gao Z., Yang H., Guo X.;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY150312; AAN73220.1; --
 DR HSP; P13811; IMEM.
 DR MEROPS; C31.001; --
 DR MEROPS; S32.001; --
 DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
 DR GO; GO:0016032; P:viral life cycle; IEA.
 DR GO; GO:0019082; P:viral protein processing; IEA.
 DR InterPro; IPR008741; Peptidase_C31.
 DR InterPro; IPR008743; Peptidase_C32.
 DR InterPro; IPR008760; Peptidase_S32.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR InterPro; IPR004008; Reg_Chr_Condens.
 DR InterPro; IPR001412; tRNA-synt 1.
 DR Pfam; PF05410; Peptidase_C31; 1.
 DR Pfam; PF05411; Peptidase_C32; 1.
 DR Pfam; PF05412; Peptidase_C33; 1.
 DR Pfam; PF05579; Peptidase_S32; 1.
 DR PROSITE; PS00178; AA_TRNA_LIGASE_1; UNKNOWN_1.
 DR PROSITE; PS00626; RCC1_2; UNKNOWN_1.
 SQ SEQUENCE 2503 AA; 271840 MW; 335ED690575FF4BE CRC64;
 Query Match 38.0%; Score 52; DB 2; Length 2503;
 Best Local Similarity 47.4%; Pred. No. 1.3e+02;
 Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 QY 1 RPYSCGAVWTLNRAIGHR 19
 DB 578 RPYSPAAATWTVSQSYARH 596
 RESULT 5
 Q7W695 PRELIMINARY; PRT; 365 AA.
 ID Q7W695
 AC Q7W695;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Fimbrial adhesin.
 GN Name=fimb; Synonyms=fhaB; OrderedLocusNames=BFP3023;
 OS Bordetella parapertussis.

Matches 14; Conservative 1; Indels 5; Mismatches 18; Gaps 2;

QY 1 RPVSFCG---AVWT-----LNRAIGRHF 20
 :||||| |||
Db 15 RVISFGAALAWAGLAVQFAMVDPVDCGRAIGLHF 52

RESULT 7
Q69418 ID Q69418 PRELIMINARY; PRT; 1693 AA.
AC Q69418;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE ORF1, ORF2 & ORF3.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]

SEQUENCE FROM N.A.
RP von Brunn A., Seebach J., Thyagarajan S.P., Mohanavalli B., Menon T.,
RA Procsner G.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RR EXBL: X99441; CAA67802.1; -;
DR GO: GO:0008174; F:RNA methyltransferase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003724; F:RNA helicase activity; IEA.
DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0006396; P:RNA processing; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR GO: GO:0019079; P:viral genome replication; IEA.
DR GO: GO:0016032; P:vital life cycle; IEA.
DR GO: GO:0019082; P:viral protein processing; IEA.
DR InterPro: IPR002589; Alpp.
DR InterPro: IPR008748; Peptidase_C41.
DR InterPro: IPR001788; RNA dep RNAPol2.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR InterPro: IPR000606; Viral_helicase.
DR InterPro: IPR002588; V_methyltrans.
DR Pfam: PF01661; Alpp; 1.
DR Pfam: PF05417; Peptidase_C41; 1.
DR Pfam: PF00978; RNA dep RNAPol2; 1.
DR Pfam: PF01443; Viral_helicase1; 1.
DR Pfam: PF01660; Vmethyltransf; 1.
DR SMART; SMK0506; Alpp; 1.
SQ SEQUENCE 1693 AA; 185949 MW; 489B4ADBA5E7E529 CRC64;

Query Match 37.2%; Score 51; DB 2; Length 1693;
Best Local Similarity 38.1%; Pred.No. 1.3e+02;
Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 4 SCGAVWTLNRAIGRHFRGS 24
 :||:::||:||||:
Db 647 AFCSALYRFNRVQRHSLTGN 667

RESULT 8
Q35802 ID Q35802 PRELIMINARY; PRT; 932 AA.
AC Q35802;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Inter-alpha-inhibitor H4 heavy chain.
GN Name=ITIH4;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.

DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Non-structural proteins.
 OS Hepatitis E virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;
 OC Hepatitis E-like viruses.
 CX NCBI_taxid=12461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22811791; PubMed=12931035;
 RA Liu Z., Chi B., Takahashi K., Mishihiro S.;
 RT "A genotype IV hepatitis E virus strain that may be indigenous to
 ET Changchun, China.";
 RL Intervirology 46:252-255(2003).
 DR EMBL; AB108537; BAC71767.1; -;
 DR GO; GO:0008174; F:RNA binding; IEA.
 DR GO; GO:0003723; F:RNA helicase activity; IEA.
 DR GO; GO:0003724; F:RNA helicase activity; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0006395; P:RNA processing; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0016032; P:viral life cycle; IEA.
 DR GO; GO:0019082; P:viral protein processing; IEA.
 DR InterPro; IPR002589; Alpp.
 DR InterPro; IPR008748; Peptidase C41.
 DR InterPro; IPR001788; RNA_dep_RNapol2.
 DR InterPro; IPR007095; RNA_pol_PS.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR006066; Viral_helicase1.
 DR InterPro; IPR002588; V_methyltrans.
 DR Pfam; PF01661; Alpp; 1.
 DR Pfam; PF05417; Peptidase C41; 1.
 DR Pfam; PF00978; RNA_dep_RNapol2; 1.
 DR Pfam; PF01443; Viral_helicase1; 1.
 DR Pfam; PF01660; Vmethyltransf; 1.
 DR SMART; SM00506; Alpp; 1.
 SQ SEQUENCE 1706 AA; 187032 MW; 4980167080EAD21C CRC64;

 Query Match 36.5%; Score 50; DB 2; Length 1706;
 Best Local Similarity 45.0%; Pred. No. 1.9e+02;
 Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

 QY 4 SFGGAVTTLNRAIGRHFVRG 23
 Db 647 AFCSALYFNRICQIRHSLIG 666

 RESULT 12
 Q9V516
 ID Q9V516 PRELIMINARY; PRT; 573 AA.
 AC
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE CG8172-PA.
 GN ORFNames=CG8172;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 CX NCBI_taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA Sutton G.G., Wortman J.R., Richards S., Ashburner M., Henderson S.N.,
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle E.G., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fodor C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
 RA Pacel J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
 RT melanogaster euchromatic genome sequence.";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 RT a genomics perspective.";
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426089; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tuzy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Battencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.C.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX FLYBASE;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RX FLYBASE;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 RL -1- SIMILARITY: Belongs to peptidase family S1.

```

DR EMBL; AE003835; AAF59006.1; -.
DR HSSP; P00750; IRTF.
DR FlyBase; Fgn0033362; CG8172.
DR GO; GO:0004295; P:trypsin activity; NAS.
DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 573 AA; 62084 MW; ECDA96CD97BEE0 CRC64;

Query Match 36.1%; Score 49.5; DB 2; Length 573;
Best Local Similarity 55.0%; Pred. No. 77;
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 RPVSFGAVWT-LNRAIGH 19
DB 285 RPVPGCGEYVYTRSNRIVG 304

RESULT 13
P93714
ID P93714 PRELIMINARY; PRT; 341 AA.
AC P93714;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE PETH; ZPT2-6.
OS Petunia hybrida (Petunia).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Petunia.
OX NCBI_TaxID=4102;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Stamen;
RA Kobayashi A., Sakamoto A., Kubo K., Rybka Z., Kanno Y., Takatsuji H.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB000452; BAA19111.1; -.
DR HSSP; Q38895; INJQ.
DR TRANSFAC; T03943; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR007087; ZnF_C2H2.
DR Pfam; PF00096; zf-C2H2; 2.
DR SMART; SM00355; ZnF_C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 2.
SQ SEQUENCE 341 AA; 38431 MW; 4122B0C61505D1F4 CRC64;

Query Match 35.8%; Score 49; DB 2; Length 341;
Best Local Similarity 47.4%; Pred. No. 56;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 4 SFGCAVWTLNRAIGHFVR 22
DB 243 SVCGAFTSGQALGHWR 261

RESULT 14
Q7WI64
ID Q7WI64 PRELIMINARY; PRT; 365 AA.
AC Q7WI64;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Fimbrial adhesin.
GN Name=fimD; Synonyms=fhaE; OrderedLocusNames=BB3989;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RC MEDLINE=22827954; PubMed=101038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrrell B.G., Masekell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640446; CAE33481.1; -.
DR GO; GO:0009289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
KW Complete proteome.
SQ SEQUENCE 365 AA; 39069 MW; 25199A15E2127A2C CRC64;

Query Match 35.8%; Score 49; DB 2; Length 365;
Best Local Similarity 34.2%; Pred. No. 59;
Matches 13; Conservative 2; Mismatches 5; Indels 18; Gaps 2;

QY 1 RPVSFGAVWT-LNRAIGHF 20
DB 4 RVISFGAALVWAGLVQFAMVDPVDCGALGHF 41

RESULT 15
FHA_E BORPE
ID FHA_E BORPE STANDARD; PRT; 376 AA.
AC Q00879;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Protein fhaE precursor.
GN Name=fhaE; Synonyms=fimD; OrderedLocusNames=BP1883;
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RC MEDLINE=92371423; PubMed=1354611;
RA Loch C., Geoffroy M.C., Renaud G.;
RT "Common accessory genes for the Bordetella pertussis filamentous
RT hemagglutinin and fimbriae share sequence similarities with the papC
RT and papD gene families.";
RL EMBO J. 11:3175-3183(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=94018656; PubMed=8105363;
RA Willens R.J.L., Geuijen C., van der Heide H.G.J., Matheson M.,
RA Robinson A., Versluis L.F., Ebberink R., Theelen J., Mooi F.R.;
RT "Isolation of a putative fimbrial adhesin from Bordetella pertussis
RT and the identification of its gene.";
RL Mol. Microbiol. 9:623-634(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;

```

```

RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X66729; CAA47267.1; -.
DR EMBL; X64876; CAA46091.1; -.
DR EMBL; BX640416; CAE42166.1; ALT_INIT.
DR PIR; S36247; S36247.
DR InterPro; IPR008966; Adhes_bact.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
KW Complete proteome; Signal.
FT SIGNAL 1 37 Potential.
FT CHAIN 38 376 Protein fhaE.
SQ SEQUENCE 376 AA; 40528 MW; 2820D28FCBD4ECF CRC64;
Query Match 35.8%; Score 49; DB 1; Length 376;
Best Local Similarity 34.2%; Pred No. 61;
Matches 13; Conservative 2; Mismatches 5; Indels 18; Gaps 2;
OY 1 RPVSFCG---AVWT-----LNRAIGRHF 20
Db 15 RVISFCGAALAVWAGLVQPMADVDPVDCGRALGLHF 52

```

Search completed: November 10, 2004, 14:50:24
Job time : 35.3019 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 10.6151 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-58

Perfect score: 144

Sequence: 1 HAVARLLHIGAINFQRLDFIEQLSAPPA 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCUTS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	53	36.8	441	4	US-09-270-767-47360
2	46.5	32.3	184	4	US-10-140-002-76
3	46	31.9	106	4	US-09-270-767-58137
4	46	31.9	552	4	US-09-801-191A-2
5	46	31.9	552	4	US-09-801-191A-6
6	46	31.9	576	4	US-09-801-191A-4
7	46	31.9	576	4	US-09-801-191A-5
8	45.5	31.6	1233	4	US-09-328-352-5959
9	45	31.2	271	4	US-09-252-991A-24319
10	45	31.2	773	4	US-09-252-991A-31403
11	45	31.2	1132	2	US-08-567-508C-2
12	45	31.2	1132	3	US-09-196-480-2
13	45	31.2	1132	4	US-09-972-800A-16
14	45	31.2	1139	3	US-09-046-158A-22
15	44	30.6	87	4	US-09-543-681A-6899
16	44	30.6	147	4	US-09-248-796A-27890
17	44	30.6	220	3	US-08-297-431B-31
18	44	30.6	221	3	US-08-297-431B-2
19	44	30.6	221	3	US-08-297-431B-4
20	44	30.6	221	3	US-08-297-431B-6
21	44	30.6	221	3	US-08-297-431B-8
22	44	30.6	221	3	US-08-297-431B-10
23	44	30.6	221	3	US-08-297-431B-12
24	44	30.6	221	3	US-08-297-431B-14
25	44	30.6	221	3	US-08-297-431B-16
26	44	30.6	221	3	US-08-297-431B-18
27	44	30.6	221	3	US-08-297-431B-20

28	43.5	30.2	379	4	US-09-491-577-6	Sequence 6, Appli
29	43.5	30.2	450	4	US-09-252-991A-31528	Sequence 31528, A
30	43	29.9	445	4	US-09-252-991A-26149	Sequence 26149, A
31	43	29.9	1722	4	US-09-538-092-1033	Sequence 1033, Ap
32	42.5	29.5	235	4	US-09-270-767-41614	Sequence 41614, A
33	42.5	29.5	1174	4	US-09-252-991A-29279	Sequence 29279, A
34	42	29.2	220	4	US-09-328-352-5959	Sequence 5959, Ap
35	42	29.2	285	4	US-09-252-991A-19996	Sequence 19996, A
36	42	29.2	329	4	US-09-252-991A-30436	Sequence 30436, A
37	42	29.2	355	3	US-09-134-001C-3580	Sequence 3580, Ap
38	42	29.2	510	4	US-09-270-767-46399	Sequence 46399, A
39	42	29.2	550	4	US-09-252-991A-17682	Sequence 17682, A
40	42	29.2	628	2	US-08-394-177-6	Sequence 6, Appli
41	42	29.2	628	3	US-08-394-912A-6	Sequence 6, Appli
42	42	29.2	628	4	US-09-333-636-6	Sequence 6, Appli
43	42	29.2	711	4	US-09-252-991A-19682	Sequence 19682, A
44	42	29.2	774	4	US-09-252-991A-19984	Sequence 19984, A
45	42	29.2	836	4	US-09-252-991A-23513	Sequence 23513, A

ALIGNMENTS

RESULT 1

US-09-270-767-47360
; Sequence 47360, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47360
; LENGTH: 441
; TYPE: PRT
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-47360

Query Match 36.8%; Score 53; DB 4; Length 441;
Best Local Similarity 47.4%; Pred. No. 2.2;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 9 HIGAINFQRLDFIEQLSAP 27
|||:::|||||
Db 37 HTESLLFRAIDLTJEQLXP 55

RESULT 2

US-10-140-002-76
; Sequence 76, Application US/10140002
; Patent No. 6725730
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330RIC59
; CURRENT APPLICATION NUMBER: US/10/140,002
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 76
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-002-76

Query Match 32.3%; Score 46.5; DB 4; Length 184;
Best Local Similarity 48.3%; Pred. No. 8.7;
Matches 14; Conservative 3; Mismatches 9; Indels 3; Gaps 1;

QY 4 VVARLLHIGAIMFOR---LDFIEQLSAPPA 29
DB 93 VSALQHLKALDLRSNQFQDFPEQLTALPA 121

RESULT 3
US-09-270-767-58137
; Sequence 58137, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58137
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-58137

Query Match 31.9%; Score 46; DB 4; Length 106;
Best Local Similarity 45.0%; Pred. No. 5.4;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 HAVVARLLHIGAIMFORLDF 20
DB 87 HLTGQNLHENAVTPEKLD 106

RESULT 4
US-09-801-191A-2
; Sequence 2, Application US/09801191A
; Patent No. 6537788
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001159
; CURRENT APPLICATION NUMBER: US/09/801,191A
; CURRENT FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Human
US-09-801-191A-2

Query Match 31.9%; Score 46; DB 4; Length 552;
Best Local Similarity 36.7%; Pred. No. 39;
Matches 11; Conservative 7; Mismatches 8; Indels 4; Gaps 1;

QY 3 VVARLLHIGAIMFORL----DFIEQLSAPP 28
DB 187 VIARILHGGWVAQQGLLHVGDIIKEVNGQP 192

RESULT 5
US-09-801-191A-6
; Sequence 6, Application US/09801191A
; Patent No. 6537788
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001159
; CURRENT APPLICATION NUMBER: US/09/801,191A
; CURRENT FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Mus Musculus
US-09-801-191A-6

Query Match 31.9%; Score 46; DB 4; Length 552;
Best Local Similarity 36.7%; Pred. No. 39;
Matches 11; Conservative 7; Mismatches 8; Indels 4; Gaps 1;

QY 3 VVARLLHIGAIMFORL----DFIEQLSAPP 28
DB 163 VIARILHGGWVAQQGLLHVGDIIKEVNGQP 192

RESULT 6
US-09-801-191A-4
; Sequence 4, Application US/09801191A
; Patent No. 6537788
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001159
; CURRENT APPLICATION NUMBER: US/09/801,191A
; CURRENT FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 576
; TYPE: PRT
; ORGANISM: Human
US-09-801-191A-4

Query Match 31.9%; Score 46; DB 4; Length 576;
Best Local Similarity 36.7%; Pred. No. 42;
Matches 11; Conservative 7; Mismatches 8; Indels 4; Gaps 1;

QY 3 VVARLLHIGAIMFORL----DFIEQLSAPP 28
DB 187 VIARILHGGWVAQQGLLHVGDIIKEVNGQP 216

RESULT 7
US-09-801-191A-5
; Sequence 5, Application US/09801191A
; Patent No. 6537788
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001159
; CURRENT APPLICATION NUMBER: US/09/801,191A

```

Db 189 VARLLAVGLVAFEIVD 205

Best local similarity 30.7%; local no. 12;
Matches 11; Conservative 7; Mismatches 8; Indels 4; Gaps 1;

CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-03-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142

```

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31403
; LENGTH: 773
; TYPE: PR1
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31403

```

Query Match 31.2%; Score 45; DB 4; Length 773;
Best Local Similarity 31.2%; Pred. No. 86;
Matches 5; Conservative 9; Mismatches 2; Indels

```
QY      13 IMFQRLDFIEQLSAPP 28
        ::|:::|:::|
Db      614 VVFRQVEFVDQMGGDP 629
```

; Sequence 2, Application US/08567508C
 ; Patent No. 5914393
 ; GENERAL INFORMATION:

APPLICANT: Stuart, Susan G.
TITLE OF INVENTION: A NOVEL HUMAN JAK2 KINASE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS INC.

STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304

```
,
, COMPUTER READABLE FORM:
,
, MEDIUM TYPE: Diskette
,
, COMPUTER: IBM Compatible
,
, OPERATING SYSTEM: DOS
,
, SOFTWARE: FastSQL Version 1.5
,
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/567,508C
, FILING DATE: 05-DEC-1995
, CLASSIFICATION: 435
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER:
```

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-00491US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-843-4186
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1132 amino acids
TYPE: amino acid

```
Query Match      31.2%; Score 45; DB 4; Length 271;
Best Local Similarity 52.9%; Pred. No. 24;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
```

```

;
; LENGTH: 1132 amino acids
; TYPE: amino acid
;

```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
US-08-567-508C-2

Query Match          31.2%; Score 45; DB 2; Length 1132;
Best Local Similarity 25.0%; Pred. NO. 1.4e+02;
Matches 7; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY      2 AVVALLHIGAIMFORLDFIEQLSAPPA 29
Db      1032 SVASDVWSFGVLYELFTYIEKSPPA 1059

RESULT 12
US-09-196-480-2
; Sequence 2, Application US/09196480
; Patent No. 6019966
; GENERAL INFORMATION:
; APPLICANT: Coleman, Roger
; APPLICANT: Stuart, Susan G.
; TITLE OF INVENTION: A NOVEL HUMAN JAK2 KINASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,480
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/567,508
; FILING DATE: 05-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0049US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1132 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
US-09-196-480-2

Query Match          31.2%; Score 45; DB 3; Length 1132;
Best Local Similarity 25.0%; Pred. NO. 1.4e+02;
Matches 7; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY      2 AVVALLHIGAIMFORLDFIEQLSAPPA 29
Db      1032 SVASDVWSFGVLYELFTYIEKSPPA 1059
```

```
RESULT 13
US-09-972-800A-16
; Sequence 16, Application US/09972800A
; Patent No. 6534277
; GENERAL INFORMATION:
; APPLICANT: Hancock, W. E.
; APPLICANT: Okaynak, E.
; TITLE OF INVENTION: ROLES OF JAK/STAT FAMILY MEMBERS IN TOLERANCE
; FILE REFERENCES: 7853-192
; CURRENT APPLICATION NUMBER: US/09/972,800A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US/09/549,654
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 1132
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-800A-16

Query Match          31.2%; Score 45; DB 4; Length 1132;
Best Local Similarity 25.0%; Pred. NO. 1.4e+02;
Matches 7; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY      2 AVVALLHIGAIMFORLDFIEQLSAPPA 29
Db      1032 SVASDVWSFGVLYELFTYIEKSPPA 1059

RESULT 14
US-09-046-158A-22
; Sequence 22, Application US/09046158A
; Patent No. 6187552
; GENERAL INFORMATION:
; APPLICANT: Roberts, Steven L.
; APPLICANT: Kaytes, Paul S.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING INHIBITORS OF
; TITLE OF INVENTION: JAK2/CYTOKINE RECEPTOR BINDING
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pharmacia & Upjohn Co., Intellectual Property
; ADDRESSEE: Legal Services
; STREET: 301 Henrietta Street
; CITY: Kalamazoo
; STATE: MI
; COUNTRY: USA
; ZIP: 49001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/046,158A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Darnley Jr., James D.
; REGISTRATION NUMBER: 33,673
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 616/833-2210
; TELEFAX: 616/833-8897
; TELEX: 224401
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1139 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-046-158A-22
```


Query Match 31.2%; Score 45; DB 3; Length 1139;
 Best Local Similarity 25.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 10; Mismatches 11; Indels 0; Gaps 0;
 QY 2 AVVARLLHIGAIMFQRLDFEQLSAPPA 29
 DB 1030 SVASDWSFGVVLIELFTYIEKSKSPPA 1057

RESULT 15
 US-09-543-681A-6899
 ; Sequence 6899, Application US/09543681A
 ; Patent No. 6605709
 ; GENERAL INFORMATION:
 ; APPLICANT: GARY BRETON
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 ; FILE REFERENCE: 2709.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/543,681A
 ; CURRENT FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: US 60/128,706
 ; PRIOR FILING DATE: 1999-04-09
 ; NUMBER OF SEQ ID NOS: 8344
 ; SEQ ID NO 6899
 ; LENGTH: 87
 ; TYPE: PRT
 ; ORGANISM: Proteus mirabilis
 US-09-543-681A-6899

Query Match 30.6%; Score 44; DB 4; Length 87;
 Best Local Similarity 58.3%; Pred. No. 9;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 7 LLHIGAIMFQRL 18
 DB 20 LVHLGMILFQRI 31

Search completed: November 10, 2004, 14:55:41
 Job time : 10.6651 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 33.5415 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-58

Perfect score: 144
Sequence: 1 HAVVARLLHIGAIMFORLDFIEQLSAPPA 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09E_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	144	100.0	29	US-10-092-750-58	Sequence 58, Appl
2	51.5	35.8	176	US-10-424-599-153449	Sequence 153449,
3	49	34.0	112	US-10-425-115-204154	Sequence 204154,
4	49	34.0	392	US-10-437-963-163286	Sequence 163286,
5	48	33.3	318	US-10-425-114-64372	Sequence 64372, A
6	48	33.3	1008	US-10-369-493-5629	Sequence 5629, AD
7	48	33.3	1008	US-10-369-493-5630	Sequence 5630, AD
8	47.5	33.0	325	US-10-369-493-11896	Sequence 11896, A
9	47.5	33.0	623	US-10-437-963-127647	Sequence 127647,
10	47.5	33.0	645	US-10-437-963-127646	Sequence 127646,
11	47	32.6	795	US-10-156-761-9690	Sequence 9690, Ap
12	46.5	32.3	146	US-10-276-774-1599	Sequence 1599, Ap
13	46.5	32.3	184	US-10-028-072-76	Sequence 76, Appl

14	46.5	32.3	184	14	US-10-140-809-76	Sequence 76, Appl
15	46.5	32.3	184	14	US-10-121-049-76	Sequence 76, Appl
16	46.5	32.3	184	14	US-10-123-904-76	Sequence 76, Appl
17	46.5	32.3	184	14	US-10-140-470-76	Sequence 76, Appl
18	46.5	32.3	184	14	US-10-175-746-76	Sequence 76, Appl
19	46.5	32.3	184	14	US-10-176-918-76	Sequence 76, Appl
20	46.5	32.3	184	14	US-10-176-921-76	Sequence 76, Appl
21	46.5	32.3	184	14	US-10-137-865-76	Sequence 76, Appl
22	46.5	32.3	184	14	US-10-140-474-76	Sequence 76, Appl
23	46.5	32.3	184	14	US-10-142-431-76	Sequence 76, Appl
24	46.5	32.3	184	14	US-10-143-114-76	Sequence 76, Appl
25	46.5	32.3	184	14	US-10-140-002-76	Sequence 76, Appl
26	46.5	32.3	184	14	US-10-142-419-76	Sequence 76, Appl
27	46.5	32.3	184	14	US-10-123-262-76	Sequence 76, Appl
28	46.5	32.3	184	14	US-10-142-423-76	Sequence 76, Appl
29	46.5	32.3	184	14	US-10-141-755-76	Sequence 76, Appl
30	46.5	32.3	184	14	US-10-143-032-76	Sequence 76, Appl
31	46.5	32.3	184	14	US-10-123-108-76	Sequence 76, Appl
32	46.5	32.3	184	14	US-10-123-236-76	Sequence 76, Appl
33	46.5	32.3	184	14	US-10-123-261-76	Sequence 76, Appl
34	46.5	32.3	184	14	US-10-140-921-76	Sequence 76, Appl
35	46.5	32.3	184	14	US-10-140-921-76	Sequence 76, Appl
36	46.5	32.3	184	14	US-10-121-045-76	Sequence 76, Appl
37	46.5	32.3	184	14	US-10-123-292-76	Sequence 76, Appl
38	46.5	32.3	184	14	US-10-124-813-76	Sequence 76, Appl
39	46.5	32.3	184	14	US-10-124-813-76	Sequence 76, Appl
40	46.5	32.3	184	14	US-10-140-925-76	Sequence 76, Appl
41	46.5	32.3	184	14	US-10-140-925-76	Sequence 76, Appl
42	46.5	32.3	184	14	US-10-160-498-76	Sequence 76, Appl
43	46.5	32.3	184	14	US-10-124-824-76	Sequence 76, Appl
44	46.5	32.3	184	14	US-10-127-825A-76	Sequence 76, Appl
45	46.5	32.3	184	14	US-10-127-825A-76	Sequence 76, Appl

ALIGNMENTS

RESULT 1
US-10-092-750-58
; Sequence 58, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 58
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-58

Query Match 100.0%; Score 144; DB 14; Length 29;
Best Local Similarity 100.0%; Pred. No. 48-15; 0;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HAVVARLLHIGAIMFORLDFIEQLSAPPA 29
|||||
Db 1 HAVVARLLHIGAIMFORLDFIEQLSAPPA 29
|||||

RESULT 2
US-10-424-599-153449
; Sequence 153449, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

```

; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 153449
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(176)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_109589C.1.pep
; US-10-424-599-153449

Query Match      35.8%; Score 51.5; DB 15; Length 176;
Best Local Similarity 35.7%; Pred. No. 7.9;
Matches 10; Conservative 9; Mismatches 8; Indels 1; Gaps 1;

QY 1 HAVVALLHIGAIMFQRLDFIEQLSAPP 28
| : : : : : : : : : : : : : : : :
Db 114 HPPIAQVLFYEVVLI-KLDFLERILAPP 140

RESULT 3
US-10-425-115-204154
; Sequence 204154, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 204154
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(112)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_117781C.1.pep
; US-10-425-115-204154

Query Match      34.0%; Score 49; DB 17; Length 112;
Best Local Similarity 39.3%; Pred. No. 12;
Matches 11; Conservative 3; Mismatches 14; Indels 0; Gaps 0;

QY 1 HAVVALLHIGAIMFQRLDFIEQLSAPP 28
| : : : : : : : : : : : : : : : :
Db 39 HHAVWRLHYSPPFFPTLLFLVETKPP 66

RESULT 4
US-10-437-963-163286
; Sequence 163286, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 163286
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_62297C.1.pep
; US-10-437-963-163286

Query Match      34.0%; Score 49; DB 16; Length 392;
Best Local Similarity 44.4%; Pred. No. 47;
Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 12 AIMFQRLDFIEQLSAPPA 29
| : : : : : : : : : : : : : : : :
Db 347 AVVGDKVDFAEKLSPPS 364

RESULT 5
US-10-425-114-64372
; Sequence 64372, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64372
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-236-A4_FLI.pep
; US-10-425-114-64372

Query Match      33.3%; Score 48; DB 15; Length 318;
Best Local Similarity 50.0%; Pred. No. 53;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 HAVVALLHIGAIMFQRLDF 20
| : : : : : : : : : : : : : : : :
Db 268 HAVVDELEHLGALVRHGD 287

RESULT 6
US-10-369-493-5629
; Sequence 5629, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
```

```
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5629
; LENGTH: 1008
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5629

Query Match      33.3%; Score 48; DB 14; Length 1008;
Best Local Similarity 27.0%; Pred. No. 1.9e+02;
Matches 10; Conservative 10; Mismatches 5; Indels 12; Gaps 1;

QY 3 VVARLLHIGA-----IMFQRLDFIEQLSAP 27
DB 303 LMASMLHGAKDAKKEQEFELLDEKVDFTQALQMP 339

RESULT 7
US-10-369-493-5630
; Sequence 5630, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5630
; LENGTH: 1008
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5630

Query Match      33.3%; Score 48; DB 14; Length 1008;
Best Local Similarity 27.0%; Pred. No. 1.9e+02;
Matches 10; Conservative 10; Mismatches 5; Indels 12; Gaps 1;

QY 3 VVARLLHIGA-----IMFQRLDFIEQLSAP 27
DB 303 LMASMLHGAKDAKKEQEFELLDEKVDFTQALQMP 339

RESULT 8
US-10-369-493-11896
; Sequence 11896, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
```

```
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11896
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mesorhizobium loti
US-10-369-493-11896

Query Match      33.0%; Score 47.5; DB 14; Length 325;
Best Local Similarity 40.0%; Pred. No. 66;
Matches 10; Conservative 8; Mismatches 4; Indels 3; Gaps 1;

QY 5 ARLHIGAIMFQRLDFIEQLSAPPA 29
DB 3 ARLVHIGSAV---VDVYVRDALPA 24

RESULT 9
US-10-437-963-127647
; Sequence 127647, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 127647
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT MRT4530_30078C.1.pep
US-10-437-963-127647

Query Match      33.0%; Score 47.5; DB 16; Length 623;
Best Local Similarity 46.7%; Pred. No. 1.3e+02;
Matches 14; Conservative 4; Mismatches 7; Indels 5; Gaps 2;

QY 1 HAVVARLLH--IGAIMFQRLDFIEQLSAPP 28
DB 189 HGRVAKLLRHITIGASV---VDFAYQLATPP 215

RESULT 10
US-10-437-963-127646
; Sequence 127646, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 127646
```

```

US-10-276-774-1599
Query Match      32.3%; Score 46.5; DB 15; Length 146;
Best Local Similarity 48.3%; Pred. No. 39;
Matches 14; Conservative 3; Mismatches 9; Indels 3; Gaps 1;

US-10-437-963-127646
Query Match      33.0%; Score 47.5; DB 16; Length 645;
Best Local Similarity 46.7%; Pred. No. 1.4e+02;
Matches 14; Conservative 4; Mismatches 7; Indels 5; Gaps 2;

US-10-156-761-9690
Query Match      32.6%; Score 47; DB 14; Length 795;
Best Local Similarity 41.4%; Pred. No. 2.1e+02;
Matches 12; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

US-10-156-761-9690
Query Match      32.6%; Score 47; DB 14; Length 795;
Best Local Similarity 41.4%; Pred. No. 2.1e+02;
Matches 12; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

US-10-276-774-1599
Query Match      32.3%; Score 46.5; DB 15; Length 146;
Best Local Similarity 48.3%; Pred. No. 39;
Matches 14; Conservative 3; Mismatches 9; Indels 3; Gaps 1;

```

; PRIOR APPLICATION NUMBER: 60/063127
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/063327
 ; PRIOR FILING DATE: 1997-10-27
 ; PRIOR APPLICATION NUMBER: 60/063329
 ; PRIOR FILING DATE: 1997-10-27
 ; PRIOR APPLICATION NUMBER: 60/063550
 ; PRIOR FILING DATE: 1997-10-28
 ; PRIOR APPLICATION NUMBER: 60/063561
 ; PRIOR FILING DATE: 1997-10-28
 ; PRIOR APPLICATION NUMBER: 60/063704
 ; PRIOR FILING DATE: 1997-10-29
 ; PRIOR APPLICATION NUMBER: 60/063733
 ; PRIOR FILING DATE: 1997-10-29
 ; PRIOR APPLICATION NUMBER: 60/063735
 ; PRIOR FILING DATE: 1997-10-29
 ; PRIOR APPLICATION NUMBER: 60/063738
 ; PRIOR FILING DATE: 1997-10-29
 ; PRIOR APPLICATION NUMBER: 60/063755
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/064248
 ; PRIOR FILING DATE: 1997-11-03
 ; PRIOR APPLICATION NUMBER: 60/064809
 ; PRIOR FILING DATE: 1997-11-07
 ; PRIOR APPLICATION NUMBER: 60/065186
 ; PRIOR FILING DATE: 1997-11-12
 ; PRIOR APPLICATION NUMBER: 60/065846
 ; PRIOR FILING DATE: 1997-11-17
 ; PRIOR APPLICATION NUMBER: 60/066364
 ; PRIOR FILING DATE: 1997-11-21
 ; PRIOR APPLICATION NUMBER: 60/066453
 ; PRIOR FILING DATE: 1997-11-24
 ; PRIOR APPLICATION NUMBER: 60/066511
 ; PRIOR FILING DATE: 1997-11-24
 ; PRIOR APPLICATION NUMBER: 60/066770
 ; PRIOR FILING DATE: 1997-11-24
 ; PRIOR APPLICATION NUMBER: 60/069212
 ; PRIOR FILING DATE: 1997-12-11
 ; PRIOR APPLICATION NUMBER: 60/069278
 ; PRIOR FILING DATE: 1997-12-11
 ; PRIOR APPLICATION NUMBER: 60/069334
 ; PRIOR FILING DATE: 1997-12-11
 ; PRIOR APPLICATION NUMBER: 60/069694
 ; PRIOR FILING DATE: 1997-12-15
 ; PRIOR APPLICATION NUMBER: 60/072320
 ; PRIOR FILING DATE: 1998-01-23
 ; PRIOR APPLICATION NUMBER: 60/073612
 ; PRIOR FILING DATE: 1998-02-04
 ; PRIOR APPLICATION NUMBER: 60/074086
 ; PRIOR FILING DATE: 1998-02-09
 ; PRIOR APPLICATION NUMBER: 60/074092
 ; PRIOR FILING DATE: 1998-02-09
 ; PRIOR APPLICATION NUMBER: 60/077791
 ; PRIOR FILING DATE: 1998-03-12
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/079294
 ; PRIOR FILING DATE: 1998-03-25
 ; PRIOR APPLICATION NUMBER: 60/079663
 ; PRIOR FILING DATE: 1998-02-27
 ; PRIOR APPLICATION NUMBER: 60/079728
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/080165
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/081203
 ; PRIOR FILING DATE: 1998-04-09
 ; PRIOR APPLICATION NUMBER: 60/081229
 ; PRIOR FILING DATE: 1998-04-09
 ; PRIOR APPLICATION NUMBER: 60/081695
 ; PRIOR FILING DATE: 1998-04-14
 ; PRIOR APPLICATION NUMBER: 60/081817
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/081818

; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/082999
 ; PRIOR FILING DATE: 1998-04-24
 ; PRIOR APPLICATION NUMBER: 60/083322
 ; PRIOR FILING DATE: 1998-04-28
 ; PRIOR APPLICATION NUMBER: 60/083545
 ; PRIOR FILING DATE: 1998-04-29
 ; PRIOR APPLICATION NUMBER: 60/084600
 ; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/084627
 ; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/084637
 ; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/085149
 ; PRIOR FILING DATE: 1998-05-12
 ; PRIOR APPLICATION NUMBER: 60/085323
 ; PRIOR FILING DATE: 1998-05-13
 ; PRIOR APPLICATION NUMBER: 60/085338
 ; PRIOR FILING DATE: 1998-05-13
 ; PRIOR APPLICATION NUMBER: 60/085339
 ; PRIOR FILING DATE: 1998-05-13
 ; PRIOR APPLICATION NUMBER: 60/085579
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085697
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085704
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/086414
 ; PRIOR FILING DATE: 1998-05-22
 ; PRIOR APPLICATION NUMBER: 60/086430
 ; PRIOR FILING DATE: 1998-05-22
 ; PRIOR APPLICATION NUMBER: 60/087106
 ; PRIOR FILING DATE: 1998-05-28
 ; PRIOR APPLICATION NUMBER: 60/088026
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088730
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088741
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088810
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088858
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/089532
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089599
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089907
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089947
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/090349
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090429
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090445
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090538
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090863
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/091360
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07

Query Match 32.3%; Score 46.5; DB 14; Length 184;
 Best Local Similarity 48.3%; Pred No. 50;
 Matches 14; Conservative 3; Mismatches 9; Indels 3; Gaps 1;
 QY 4 VARLLHIGAIMFOR---LDFIEQLSAPPA 29

Result No.	Query		Length	DB	ID	Description
	Score	Match				
1	50.5	35.1	274	2	T03797	thymidylate kinase
2	50	34.7	740	2	D71602	hypothetical prote
3	50	34.7	1411	2	T18417	hypothetical prote
4	50	34.7	1417	2	T18418	hypothetical prote
5	48.5	33.7	742	2	S53663	hydrogenase-relate
6	48	33.3	192	2	A10758	precorrin-8W decar
7	48	33.3	342	2	S23438	hypothetical prote
8	48	33.3	610	2	D23432	hypothetical prote
9	48	33.3	651	2	AH3276	lysothaphin (EC 3.
10	48	33.3	842	2	AB0674	probable hydrolase
11	48	33.3	1008	2	T19832	probable RNA helic
12	46	31.9	260	2	C82081	cell division prot
13	46	31.9	576	2	A57653	disks large homolo
14	45	31.2	275	2	C81700	3'-deoxy-manno-ocu
15	45	31.2	537	2	T41666	carl homolog - fis
16	45	31.2	1332	2	JW0091	Janu kinase (EC 2
17	45	31.2	1201	2	T85603	kinasin-related pr
18	44.5	30.9	513	2	S21535	nitrate reductase
19	44	30.6	206	1	I40173	crostate phosphorib
20	44	30.6	215	2	E75058	hypothetical prote
21	44	30.6	221	1	A26753	glutathione transp
22	44	30.6	303	2	T23583	hypothetical prote
23	44	30.6	345	2	H95381	probable iron upta
24	44	30.6	469	2	G70699	probable rdaA prot
25	44	30.6	540	2	S72323	transcription fact
26	44	30.6	1515	2	S51824	myosin heavy chain
27	43.5	30.2	469	2	AD1926	hypothetical prote
28	43.5	30.2	505	2	S75722	UDP-N'-acetylmuram
29	43.5	30.2	834	2	T19010	hypothetical prote

Matches	10;	Conservative	5;	Mismatches	5;	Indels	2;	Gaps	1;
---------	-----	--------------	----	------------	----	--------	----	------	----

```

QY 7 LLHIGAIMFORLDFIEQLSAPP 28
   |:::|:::|:::|:::|:::|
Db 16 LVHVS--FLQLDFHQLNEPP 35

RESULT 3
Ti8417
Hypothetical protein C0120w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18417
R:Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, July 1997
A:Reference number: Z18934
A:Accession: T18417
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1411 <LAW>
A:Cross-references: UNIPROT:O77309; EMBL:Z97348; PIDN:CAB10571.2
C:Genetics:
A:Introns: 288/1; 490/1; 498/1; 519/1; 556/1; 613/1; 666/1; 921/3
A>Note: C0120w
C:Superfamily: Plasmodium falciparum hypothetical protein C0120w

Query Match 34.7%; Score 50; DB 2; Length 1411;
Best Local Similarity 45.5%; Pred. No. 17;
Matches 10; Conservative 5; Mismatches 5; Indels 2; Gaps 1;

QY 7 LLHIGAIMFORLDFIEQLSAPP 28
   |:::|:::|:::|:::|:::|
Db 686 LVHVS--FLQLDFHQLNEPP 705

RESULT 4
Ti8418
Hypothetical protein C0120w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18418
R:Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, July 1997
A:Reference number: Z18934
A:Accession: T18418
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1417 <LAW>
A:Cross-references: UNIPROT:O77310; EMBL:Z97348; PIDN:CAB10572.2
C:Genetics:
A:Introns: 286/1; 478/1; 496/1; 517/1; 554/1; 616/1; 669/1; 924/3
A>Note: C0120w
C:Superfamily: Plasmodium falciparum hypothetical protein C0120w

Query Match 34.7%; Score 50; DB 2; Length 1417;
Best Local Similarity 45.5%; Pred. No. 17;
Matches 10; Conservative 5; Mismatches 5; Indels 2; Gaps 1;

QY 7 LLHIGAIMFORLDFIEQLSAPP 28
   |:::|:::|:::|:::|:::|
Db 689 LVHVS--FLQLDFHQLNEPP 708

RESULT 5
S53663
hydrogenase-related protein V - Azotobacter chroococcum
C:Species: Azotobacter chroococcum
C>Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C:Accession: S53663
R:Du, L.; Tidelius, K.H.; Souza, E.M.; Garg, R.P.; Yates, M.G.
J. Mol. Biol. 243, 549-557, 1994
A:title: Sequences, organization and analysis of the hupZMNOQRTV genes from the Azotobac
A:Reference number: S53655; MUID:95055698; PMID:7966281
A:Accession: S53663

QY 7 LLHIGAIMFORLDFIEQLSAPP 29
   |:::|:::|:::|:::|:::|
Db 231 HVGALLRHTAQAIQSPSPPPA 254

RESULT 6
AI0758
precorrin-8W decarboxylase (EC 1.-.-.-) [imported] - Salmonella enterica subsp. enteric
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A>Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 16-Aug-2004
C:Accession: AI0758
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher
th, T.; Conington, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AI0758
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-192 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD02391.1; PID:gi6503262; GSPDB:GN00176
C:Genetics:
A:Gene:cbiI
C:Superfamily: Precorrin-6Y methyltransferase (decarboxylating) CbiI
C:Keywords: oxidoreductase

Query Match 33.3%; Score 48; DB 2; Length 192;
Best Local Similarity 44.4%; Pred. No. 4.1;
Matches 12; Conservative 7; Mismatches 4; Indels 4; Gaps 2;

QY 1 HAVVARLLHIGAIMFORLDFIEQLS 26
   |:::|:::|:::|:::|:::|
Db 138 HSAHLVHIGAC---RMDCVQLQSS 161

RESULT 7
S23438
Hypothetical protein 3 - Azotobacter vinelandii
C:Species: Azotobacter vinelandii
C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 22-Jun-2003
C:Accession: S23438
R:Chen, J.C.; Mortenson, L.E.
Biochim. Biophys. Acta 1131, 199-202, 1992
A:title: Identification of six open reading frames from a region of the Azotobacter vin
A:Reference number: S23438; MUID:92305060; PMID:1610901
A:Accession: S23438
A:Molecule type: DNA
A:Residues: 1-342 <CHE>
A:Cross-references: EMBL:X63650
C:Genetics:
A:Start codon: CTG
C:Superfamily: hydrogenase maturation factor, HupK type

Query Match 33.3%; Score 48; DB 2; Length 342;
Best Local Similarity 38.1%; Pred. No. 7.7;
Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 9 HIGAIMFORLDFIEQLSAPP 29
   |:::|:::|:::|:::|:::|
Db 231 HVGALLRHTAQAIQSPSPPPA 251

```


A;Title: DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: C82081

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-260 <HEI>

A;Cross-references: UNIPROT:Q9KPG9; GB:AE004310; GB:AE003852; NID:g9656963; PIDN:AAF9554

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC2399

A;Map position: 1

C;Superfamily: cell division protein ftsQ

Query Match 31.9%; Score 46; DB 2; Length 260;

Best Local Similarity 42.9%; Pred. No. 12;

Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 3 VVARLHIGAIMFQRLDFIEQ 23

DB 74 VLARLDHIGTMSQDINVLQE 94

RESULT 13

AS7653

disks large homolog DLG2 - human

C;Species: Homo sapiens (man)

C;Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 09-Jul-2004

C;Accession: AS7653

R;Mazoyer, S.; Gayther, S.A.; Nagai, M.A.; Smith, S.A.; Dunning, A.; van Rensburg, E.J.;

Genomics 28, 25-31, 1995

A;Title: A gene (DLG2) located at 17q12-q21 encodes a new homologue of the Drosophila tu

A;Reference number: AS7653; MUID:96070428; PMID:7590743

A;Accession: AS7653

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-576 <MAZ>

A;Cross-references: UNIPROT:Q14168; GB:X82895; NID:g939884; PIDN:CAA58067.1; PID:g939885

C;Genetics:

A;Gene: GDB:DLG2; GDB:MPP2

A;Cross-references: GDB:6053914; GDB:567459; OMIM:600723

A;Map position: 17q12-17q21

F;170-239/Domain: GLGF domain homology <GLG>

F;263-311/Domain: SH3 homology <SH3>

F;375-564/Domain: guanylate kinase homology <GKI>

Query Match 31.9%; Score 46; DB 2; Length 576;

Best Local Similarity 36.7%; Pred. No. 29;

Matches 11; Conservative 7; Mismatches 8; Indels 4; Gaps 1;

QY 3 VVARLHIGAIMFQRL---DFIEQLSAPP 28

DB 187 VIARLHGNGVACQGLLHVGDIIKEVNGQP 216

RESULT 14

C81700

3-deoxy-manno-octulosonate cytidyltransferase TC0454 [imported] - Chlamydia muridarum

C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn

C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 18-Aug-2000

C;Accession: C81700

R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,

C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,

Nucleic Acids Res. 28, 1397-1406, 2000

A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A;Reference number: A81500; MUID:20150255; PMID:10684935

A;Accession: C81700

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-275 <TET>

A;Cross-references: GB:AE002314; GB:AE002160; NID:g7190495; PIDN:AAF39307.1; PID:g719049

C;Genetics:

A;Experimental source: strain Nigg (MoPn)

A;Gene: TC0454

C;Superfamily: 3-deoxy-manno-octulosonate cytidyltransferase

Query Match 31.2%; Score 45; DB 2; Length 275;

Best Local Similarity 40.9%; Pred. No. 19;

Matches 9; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 8 LHIGAIMFQSLDFIEQLSAPPA 29

DB 200 LHIGVAFRRAPFLSEYVKIPPS 221

RESULT 15

T41666

Carl homolog - fission yeast (*Schizosaccharomyces pombe*)

C;Species: *Schizosaccharomyces pombe*

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: T41666

R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.

submitted to the EMBL Data Library, May 1998

A;Reference number: Z22008

A;Accession: T41666

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-537 <MOO>

A;Cross-references: UNIPROT:O59833; EMBL:AL023590; PIDN:CAA19073.1; GSPDB:GN00068; SPDB

A;Experimental source: strain 972h-; cosmid c965

C;Genetics:

A;Gene: SPDB:SPCC965.13

A;Map position: 3

C;Superfamily: benomyl/methotrexate resistance protein

Query Match 31.2%; Score 45; DB 2; Length 537;

Best Local Similarity 47.1%; Pred. No. 39;

Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 AVVARLHIGAIMFQRL 18

DB 466 AIAAIVHLGIIMFDNM 482

Search completed: November 10, 2004, 14:52:26

Job time : 8.11321 secs

RESULT 2		
Q8A2U7		
ID	Q8A2U7	PRELIMINARY; PRT; 77 AA.
AC	Q8A2U7;	
DT	01-JUN-2003	(TRENBLrel. 24, Created)

```

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BT3208;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12663928;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DR EMBL; AE016939; AAO78314.1; -.
DR InterPro; IPR001452; SH3.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 77 AA; 9146 MW; 6F6D49C357A0F215 CRC64;

Query Match 35.4%; Score 51; DB 2; Length 77;
Best Local Similarity 52.6%; Pred. No. 3.7;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 3 VVARLHIGAIMFQRLDFI 21
DB 12 VVAVLLSLGALFFQLNFV 30

RESULT 3
O60970 PRELIMINARY; PRT; 274 AA.
AC O60970;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE TKRPI.
GN Name=TKRPI;
OS Leishmania major.
OC Eukaryota; Euclenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=99178987; PubMed=10077609;
RA Myler P.J., Audleman L., deVos T., Hixson G., Kiser P., Lemley C.,
RA Magness C., Rickel E., Sisk E., Sunkin S., Swartzell S., Westlake T.,
RA Bastien P., Fu G., Ivens A., Stuart K.;
RT "Leishmania major Friedlin chromosome 1 has an unusual distribution of
RT protein-coding genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:2902-2906(1999).
DR EMBL; AE001274; AAC24621.1; -.
DR PIR; A81456; T02797.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004798; F:thymidylate kinase activity; IEA.
DR GO; GO:0004233; P:tdp biosynthesis; IEA.
DR GO; GO:0008235; P:tdp biosynthesis; IEA.
DR InterPro; IPR000062; Thymidylate_kin.
DR Pfam; PF02223; Thymidylate_kin; 1.
SQ SEQUENCE 274 AA; 29782 MW; B27AE442261FF7D3 CRC64;

Query Match 35.1%; Score 50.5; DB 2; Length 274;
Best Local Similarity 44.1%; Pred. No. 17;
Matches 15; Conservative 5; Mismatches 7; Indels 7; Gaps 2;

QY 2 AVVARLHIGAIMFQRLD-----FIEQLSAPP 28
DB 5 AALARTLHGAGNVYRSLDCLKVLHVAEQSSPPP 38

RESULT 4
O77309 PRELIMINARY; PRT; 1416 AA.

```

```

ID O77309 PRELIMINARY; PRT; 1416 AA.
AC O77309; O77307;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cytoadherence linked asexual protein, CLAG.
GN Name=MAJ3P1.2; Synonyms=PFC0110w;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
[1]
RN SEQUENCE FROM N.A.
RC MEDLINE=99376085; PubMed=10448855;
RA Bowman S., Lawson D., Baaham D., Brown D., Chillingworth T.,
RA Churcher C.M., Craig A., Davies R.M., Devlin K., Feltwell T.,
RA Gentles S., Gwilliam R., Hamlin N., Harris D., Holroyd S., Hornsby T.,
RA Horrocks P., Jagels K., Jassal B., Kyes S., McLean J., Moule S.,
RA Mungall K., Murphy L., Oliver K., Quail M.A., Rajandream M.-A.,
RA Rutter S., Skelton J., Squares R., Squares S., Sulston J.E.,
RA Whitehead S., Woodward J.R., Newbold C., Barrell B.G.;
RT "The complete nucleotide sequence of chromosome 3 of Plasmodium
RT falciparum.";
RL Nature 400:532-538(1999).
RN [2]
RN SEQUENCE FROM N.A.
RC MEDLINE=22255708; PubMed=12368867;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C., Corton C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price S., Quail M.A., Rabinowitch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002).
DR EMBL; Z97348; CAB10571.3; -.
DR PIR; T18417; T18417.
DR InterPro; IPR005553; CLAG.
DR Pfam; PF03805; CLAG; 1.
SQ SEQUENCE 1416 AA; 167488 MW; 1E15EF2D1CD91BB5 CRC64;

Query Match 34.7%; Score 50; DB 2; Length 1416;
Best Local Similarity 45.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 5; Mismatches 5; Indels 2; Gaps 1;

QY 7 LLHIGAIMFQRLDFIEQLSAPP 28
DB 691 LVHVS--FLQDFFHQLNEPP 710

RESULT 5
O77310 PRELIMINARY; PRT; 1417 AA.
AC O77310; O77311;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytoadherence linked asexual protein, CLAG.
GN Name=MAJ3P1.5; Synonyms=PFC0120w;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
[1]
RN SEQUENCE FROM N.A.
RC MEDLINE=99376085; PubMed=10448855;

```

```

Bowman S., Lawson D., Basham D., Brown D., Chillingworth T.,
Churcher C.M., Craig A., Davies R.M., Devlin K., Felwell T.,
Gentles C., Gwilliam R., Hamlin N., Harris D., Holroyd S., Hornsby T.,
Horrocks P., Jagels K., Jassal K., Kyes S., McLean J., Moule S.,
Mungall K., Murphy L., Oliver K., Quail M.A., Rajandream M.-A.,
Rutter S., Skelton J., Squares R., Squares S., Sulston J.E.,
Whitehead S., Woodward J.R., Newbold C., Barrell B.G.;
"The complete nucleotide sequence of chromosome 3 of Plasmodium
falci-parum.";
RL Nature 400:532-538 (1999).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=22255708; PubMed=12368867;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
Mungali K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
Bucklee C.O., Burrows C., Cheruvach I., Chillingworth C.,
Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
Felwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
Line A., Maddison M., Mclean J., Mooney P., Moule S., Murphy L.,
Oliver K., Ormond D., Price C., Quail M.A., Rabbinowitsch E.,
Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
Sulston J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmodium falci-parum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002).
DR EMBL; Z97348; CAB10572.2; -.
DR PIR; T18418; T18418.
DR InterPro; IPR005553; CLAG.
DR Pfam; PF03805; CLAG; 1.
SQ SEQUENCE 1417 AA; 167241 MW; AE7AA0E54B5B1130 CRC64;

Query Match 34.7%; Score 50; DB 2; Length 1417;
Best Local Similarity 45.5%; Pred.No. 1.1e+02;
Matches 10; Conservative 5; Mismatches 5; Indels 2; Gaps 1;

QY 7 LLHGCAIMFQRDLFIQLSAPP 28
Db 689 LVHVS--FLQDFHQINSEPP 708

RESULT 6
ID O96279 PRELIMINARY; PRT; 1440 AA.
AC O96279; O96280;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytoadherence linked asexual protein 2.
GN Name=PF0935W;
OS Plasmodium falci-parum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxId=36329;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=95021743; PubMed=9804551;
RA Gardner M.J., Tettelin H., Carucci D.J., Cummings L.M., Aravind L.,
RA Koornik E.V., Shallow S., Mason T., Yu K., Fujii C., Pederson J.,
RA Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Portea M.,
RA Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
falci-parum.";
RL Science 282:1126-1132(1998).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
```

[illegible]

DT 02-MAR-2004 (T-EMBLrel. 27, Last annotation update)
 DE Cytochrome P-450-like protein (Fragment).
 GN CYP82.
 OS Vitis vinifera (Grape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Vitaceae; Vitis.
 OX NCBI_TaxID=29760;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Chardonnay;
 RA Bezier A., Lambert B., Baillieu F.;
 RT "Molecular cloning of partial cytochrome P450-like protein mRNA from
 RT grapevines."
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR ENBL: AY226829; AAP49697.1; -
 FT NON TER 1 1
 SQ SEQUENCE 259 AA; 28776 MW; 261C3D9532355817 CRC64;

Query Match 34.4%; Score 49.5; DB 2; Length 259;
 Best Local Similarity 41.2%; Pred. No. 23;
 Matches 14; Conservative 5; Mismatches 10; Indels 5; Gaps 2;

QY 1 HAVVARELLH---IGATMFQRLDFIEQ--LSAPPA 29
 DB 205 HLTALRLHGFELGAVDSFVDMTESPGLTAPKA 238

RESULT 9
 Q9FW86 PRELIMINARY; PRT; 203 AA.
 AC Q9FW86 Q7XC2;
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
 DE Putative cytochrome P450.
 GN ORFNames=OSJNBA0026L12.11;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoideae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
 RA Hsiao J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
 RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
 RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Buell R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA "The Rice Chromosome 10 Sequencing Consortium;
 RT "In-depth view of structure, activity, and evolution of rice
 RT chromosome 10.";
 RL Science 300:1566-1569 (2003).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
 DR ENBL: AC068924; AG13497.1; -
 DR EMBL: AE017112; AAP54605.1; -
 DR HSSP: P11712; LOG2.
 DR Gramene; Q7XC2; -
 DR Gramene; Q9FW86; -
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR InterPro; IPR002401; EP450I.

DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00463; EP450I.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 203 AA; 22990 MW; 859FFD62EB372645 CRC64;

Query Match 34.0%; Score 49; DB 2; Length 203;
 Best Local Similarity 44.4%; Pred. No. 21;
 Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 12 AIMFQRLDFIEQLSAPPA 29
 DB 158 AVGDKVDFAEKLSPPS 175

RESULT 10
 Q8PCD6 PRELIMINARY; PRT; 402 AA.
 ID Q8PCD6
 AC Q8PCD6;
 DT 01-OCT-2002 (T-EMBLrel. 22, Created)
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE 2-octaprenyl-6-methoxyphenol hydroxylase.
 GN Name=ubiH; OrderedLocustNames=XCC0798;
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=340;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33913 / NCPPB 528;
 RX MEDLINE=22022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Catarotte G., Cannavan F., Cardozo J., Chamberg F., Clapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.P.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities";
 RL Nature 417:459-463 (2002).
 DR ENBL; AE012179; AAM40113.1; -
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0006725; P:aromatic compound metabolism; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR000733; Flav_monooxygenase.
 DR InterPro; IPR000205; NAD_ES.
 DR InterPro; IPR003042; Rng_monooxygenase.
 DR InterPro; IPR010971; Ubi_OHases.
 DR InterPro; IPR011295; UbiH.
 DR Pfam; PF01360; Monooxygenase; 1.
 DR PRINTS; PR00420; RINGMONOXGNASE.
 DR TIGRfams; TIGR01988; Ubi-OHases; 1.
 DR TIGRfams; TIGR01984; UbiH; 1.
 KW Complete proteome.
 SQ SEQUENCE 402 AA; 42578 MW; AC80895FF3CCBF31 CRC64;

Query Match 34.0%; Score 49; DB 2; Length 402;
 Best Local Similarity 37.1%; Pred. No. 43;
 Matches 13; Conservative 5; Mismatches 7; Indels 10; Gaps 2;

QY 1 HAVVARELLHIGATMFQRLDFIEQ-----SAPP 28
 DB 158 AVGDKVDFAEKLSPPS 175

Db 169 HSAVRELLHIGT---EQHDFLQTLFVARVRASRP 200

RESULT 11

```

ID TRME RALSO STANDARD; PRT; 481 AA.
AC Q8V3H5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE tRNA modification GTPase trme.
GN Name=trme; Synonyms=thdP; OrderedLocusNames=RS01827;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Genin S., Attiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
CC -!- FUNCTION: Exhibits a very high intrinsic GTPase hydrolysis rate.
CC Involved in the biosynthesis of the hypermodified nucleoside 5-
CC methyaminomethyl-2-thiouridine, which is found in the wobble
CC position of some tRNAs (By similarity).
CC -!- SIMILARITY: Belongs to the era/trme family of GTP-binding
CC proteins. Trme subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL646057; CAD13533.1; -.
DR HAMAP; MF_00379; -.
DR InterPro; IPR005289; GTP-binding.
DR InterPro; IPR006073; GTP1_OBG.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR004520; ThdP.
DR PRINIS; PR00326; GTP1OBG.
DR PRINTS; PR00449; RASTRNSFRMG.
DR TIGRFAMs; TIGR00650; MG442; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR TIGRFAMs; TIGR00450; thdP; 1.
KW Complete proteome; GTP-binding; tRNA processing.
FT NP_BIND 245 262 GTP (Potential).
FT NP_BIND 292 296 GTP (Potential).
FT NP_BIND 361 364 GTP (Potential).
SQ SEQUENCE 481 AA; 50908 MW; 0D54949EE2E7B54D CRC64;

```

Query Match 34.0%; Score 49; DB 1; Length 481;
 Best Local Similarity 39.1%; Pred. No. 52;
 Matches 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 1 HAVRELLHIGALMFORLDFIEQ 23

Db 179 HALVERVILRLMLVEATLDFPEE 201

RESULT 12

Q7XQG9
 ID Q7XQG9 PRELIMINARY; PRT; 804 AA.

```

Q7XQG9;
AC Q7XQG9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE QJ000114.01.8 protein.
GN Name=QJ000114.01.8;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12447439;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Yang K., Yu S., Tang Y.,
RA Weng Q., Zhang L., Lu Y., Mu Y., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
DR EMBL; AL606452; CAE03127.3; -.
DR Gramene; Q7XQG9; -.
DR InterPro; IPR008862; Tcpl1.
DR Pfam; PF05794; Tcpl1; 1.
SQ SEQUENCE 804 AA; 87004 MW; 71A902FF3BAEA150 CRC64;

Query Match 34.0%; Score 49; DB 2; Length 804;
Best Local Similarity 52.8%; Pred. No. 89;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 9 HIGAIMFORLDFIEQLSAP 27
Db 418 YLGIWHYSLDMIRKLSAP 436

RESULT 13
Q8WRY3
ID Q8WRY3 PRELIMINARY; PRT; 1296 AA.
AC Q8WRY3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytoadherence linked asexual protein-like protein.
OS Plasmodium chabaudi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5825;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AS;
RX MEDLINE=11738716; PubMed=11738716;
RA Holt D.C., Fischer K., Tchavtchitch M., Wilson D.W., Hauquitz N.E.,
RA Hawthorne P.L., Gardiner D.L., Trenholme K.R., Kemp D.J.;
RT "Clags in P. falciparum and other species of Plasmodium.";
RL Mol. Biochem. Parasitol. 118:259-263(2001).
DR EMBL; AF387740; AAU57302.1; -.
DR InterPro; IPR005553; CLAG.
DR InterPro; IPR001202; WW_Rps5_WWP.
DR Pfam; PF03805; CLAG; 1.
DR PROSITE; PS01159; WW DOMAIN 1; UNKNOWN 1.
SQ SEQUENCE 1296 AA; 152409 MW; 35B134DF9A54EDAD CRC64;

Query Match 34.0%; Score 49; DB 2; Length 1296;
Best Local Similarity 38.5%; Pred. No. 1.5e+02;
Matches 10; Conservative 6; Mismatches 8; Indels 2; Gaps 1;

QY 3 VVARELLHIGAIMFORLDFIEQLSAPP 28

```

Db 654 VQYVIHVS--FMQLDFFHALNEPP 677

RESULT 14

Q88EQ7 PRELIMINARY; PRT; 155 AA.

AC Q88EQ7; (TREMELrel. 24, Created)

DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)

DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)

DE Hypothetical protein.

GN OrderedLocusNames=PP4396;

OS Pseudomonas putida (strain KT2440).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Pseudomonas.

OX NCBI_TaxID=160488;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22423060; PubMed=12534453;

RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,

RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,

RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,

RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,

RA Hance I., Chris Lee P., Holtzaple E.K., Scanlan D., Tran K.,

RA Moazzez A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,

RA Wedler H., Lauber J., Stjepandic D., Hohnselt J., Straetz M., Heim S.,

RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoft A., Tuemmler B.,

RA Fraser C.M.;

RT "Complete genome sequence and comparative analysis of the

RT metabolically versatile Pseudomonas putida KT2440.";

RL Environ. Microbiol. 4:799-808(2002).

DR EMEL; AS016790; AN69574.1; -.

DR TTGR; PF4396; -.

DR GO; GO:0019861; C:flagellum; IEA.

DR GO; GO:0009296; P:flagellum biogenesis; IEA.

DR InterPro; IPR007809; FlgN.

DR Pfam; PF05130; FlgN; 1.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 155 AA; 17170 MW; 9070C76F8C705C47 CRC64;

Query Match 33.7%; Score 48.5; DB 2; Length 155;

Best Local Similarity 52.2%; Pred. No. 19;

Matches 12; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

QY 2 AVVARLLHICAIMFORLDFIEQL 24

DB 78 AVAAQSPH-GDVMLOKLDMLSQL 99

RESULT 15

Q6N7G4 PRELIMINARY; PRT; 198 AA.

AC Q6N7G4;

DT 05-JUL-2004 (TREMELrel. 27, Created)

DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)

DE Hypothetical protein.

GN OrderedLocusNames=RPA2293;

OS Rhodopseudomonas palustris.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Bradyrhizobiaceae; Rhodopseudomonas.

OX NCBI_TaxID=1076;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CGA009 / ATCC BAA-98;

RX PubMed=14704707; DOI=10.1038/nbr923;

RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,

RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,

RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,

RA Harrison F.H., Gibson J., Harwood C.S.;

RT "Complete genome sequence of the metabolically versatile

RT photosynthetic bacterium Rhodopseudomonas palustris.";

RL Nat. Biotechnol. 22:55-61(2004).

DR EMEL; BX572600; CAE27734.1; -.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 198 AA; 22169 MW; DCDB2421AE73DB98 CRC64;

Query Match 33.7%; Score 48.5; DB 2; Length 198;

Best Local Similarity 48.3%; Pred. No. 25;

Matches 14; Conservative 2; Mismatches 10; Indels 3; Gaps 1;

QY 1 HAVVARLLHICAIMFORLDFIEQLSAPPA 29

DB 44 HRAVAVELLISRTIFAERLD---ALSAHPS 69

Search completed: November 10, 2004, 14:50:26

Job time : 40.6302 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 11.3472 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-59

Perfect score: 178
Sequence: 1 GCGTWSGMEAWLATVLKALPWHPTYOLEP 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pdp:*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pdp:*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB.pdp:*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB.pdp:*
- 5: /cgn2_6/ptodata/1/iaa/PTUS COMB.pdp:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pdp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51.5	28.9	644	US-09-107-532A-7277	Sequence 7277, Ap
2	51	28.7	441	US-09-543-681A-7207	Sequence 7207, Ap
3	50.5	28.4	57	US-09-513-999C-6363	Sequence 6363, Ap
4	50.5	28.4	60	US-09-270-767-34498	Sequence 34498, A
5	50.5	28.4	60	US-09-270-767-49715	Sequence 49715, A
6	50	28.1	308	US-09-489-039A-14278	Sequence 14278, A
7	49.5	27.8	502	US-09-724-797-90	Sequence 90, Appl
8	49	27.5	413	US-08-282-197C-49	Sequence 49, Appl
9	49	27.5	446	US-09-543-681A-5864	Sequence 5864, Ap
10	49	27.5	557	US-09-413-814-34	Sequence 34, Appl
11	48.5	27.2	345	US-09-489-039A-7417	Sequence 7417, Ap
12	48.5	27.2	1239	US-09-252-991A-30198	Sequence 30198, A
13	48	27.0	208	US-09-248-796A-24595	Sequence 24595, A
14	48	27.0	326	US-09-489-039A-13604	Sequence 13604, A
15	47.5	26.7	80	US-09-252-991A-22806	Sequence 22806, A
16	47	26.4	29	US-08-293-778-6	Sequence 6, Appl
17	47	26.4	138	US-07-857-224B-112	Sequence 112, App
18	47	26.4	151	US-08-614-935-28	Sequence 28, Appl
19	47	26.4	151	US-09-130-287-28	Sequence 28, Appl
20	47	26.4	211	US-09-205-258-1159	Sequence 1159, Ap
21	47	26.4	217	US-09-205-258-1162	Sequence 1162, Ap
22	47	26.4	243	US-09-541-759-7	Sequence 7, Appl
23	47	26.4	259	US-08-944-483-52	Sequence 52, Appl
24	47	26.4	259	US-10-165-442-2	Sequence 2, Appl
25	47	26.4	259	US-10-165-442-2	Sequence 4, Appl
26	47	26.4	295	US-08-338-368-2	Sequence 2, Appl
27	47	26.4	295	US-10-165-442-1	Sequence 1, Appl

28 47 26.4 295 4 US-10-165-442-3 Sequence 3, Appl
29 47 26.4 376 2 US-08-558-269-10 Sequence 10, Appl
30 47 26.4 376 3 US-09-410-882-10 Sequence 10, Appl
31 47 26.4 416 4 US-09-489-039A-13142 Sequence 13142, A
32 47 26.4 579 1 US-08-295-411-4 Sequence 4, Appl
33 47 26.4 579 2 US-08-955-471-4 Sequence 4, Appl
34 47 26.4 579 3 US-09-117-708-14 Sequence 14, Appl
35 47 26.4 579 5 PCT-US92-10242-4 Sequence 4, Appl
36 47 26.4 615 1 US-07-998-972A-3 Sequence 3, Appl
37 47 26.4 615 1 US-08-463-953-3 Sequence 3, Appl
38 47 26.4 615 5 PCT-US92-11357-3 Sequence 3, Appl
39 47 26.4 622 3 US-08-952-967-8 Sequence 8, Appl
40 47 26.4 622 4 US-09-054-272-42 Sequence 42, Appl
41 46.5 26.1 93 4 US-09-461-325-176 Sequence 176, App
42 46.5 26.1 93 4 US-10-012-542-176 Sequence 176, App
43 46.5 26.1 93 4 US-10-115-123-176 Sequence 176, App
44 46.5 26.1 129 4 US-09-461-325-379 Sequence 379, App
45 46.5 26.1 129 4 US-09-461-325-379 Sequence 379, App

ALIGNMENTS

RESULT 1
US-09-107-532A-7277
; Sequence 7277, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESS: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 7277:
SEQUENCE CHARACTERISTICS:
LENGTH: 644 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...644
SEQUENCE DESCRIPTION: SEQ ID NO: 7277:
US-09-107-532A-7277

```

RESULT 4
US-09-270-767-34498
; Sequence 34498, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34498
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-34498

Query Match      28.4%; Score 50.5; DB 4; Length 60;
Best Local Similarity 55.6%; Pred. No. 5.7;
Matches 10; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY      14  LATVLKALPWP-TYOLE 30
      | : : : : : | : : : : :
      | : : : : : | : : : : :
DB      36  LFALLKLIPWPITYGLK 53

RESULT 5
US-09-270-767-49715
; Sequence 49715, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49715
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-49715

Query Match      28.4%; Score 50.5; DB 4; Length 60;
Best Local Similarity 55.6%; Pred. No. 5.7;
Matches 10; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY      14  LATVLKALPWP-TYOLE 30
      | : : : : : | : : : : :
      | : : : : : | : : : : :
DB      36  LFALLKLIPWPITYGLK 53

RESULT 6
US-09-489-039A-14278
; Sequence 14278, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breta et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29

```

```
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 14278
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-14278

Query Match      28.1%; Score 50; DB 4; Length 308;
Best Local Similarity 41.7%; Pred. No. 45;
Matches 10; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

QY 2 QGTLWGSMEAWLATVLKALPWP 25
Db 173 EGTWGGNLA--MLTSLIGTPWP 194

RESULT 7
US-09-724-797-90
; Sequence 90, Application US/09724797
; Patent No. 6733998
; GENERAL INFORMATION:
; APPLICANT: Jon S. THORSON
; TITLE OF INVENTION: MICROMONOSPORA ECHINOSPORA GENES
; TITLE OF INVENTION: ENCODING FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: CALICHEAMICIN AND SELF-RESISTANCE THERETO
; FILE REFERENCE: 2653-40
; CURRENT APPLICATION NUMBER: US/09/724,797
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 60/111,325
; PRIOR FILING DATE: 1998-12-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 90
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Bacteria
US-09-724-797-90

Query Match      27.8%; Score 49.5; DB 4; Length 502;
Best Local Similarity 50.0%; Pred. No. 94;
Matches 14; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

QY 4 TLWGSMEAWLATVLKALPWPHTYQLEP 31
Db 408 TLLGAGL-ALAAATVLLAPVFPHPWATWP 434

RESULT 8
US-08-282-197C-49
; Sequence 49, Application US/08282197C
; Patent No. 5871720
; GENERAL INFORMATION:
; APPLICANT: Brzezinski, Ryszard
; APPLICANT: Dery, Claude V
; APPLICANT: Beaulieu, Carole
; TITLE OF INVENTION: Thermostable Xylanase DNA, Protein and
; TITLE OF INVENTION: Methods of Use
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Ave., NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,197C
; FILING DATE: 29-JUL-1994
```

```
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1050.0410000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 413 amino acids
; TYPE: amino acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-282-197C-49

Query Match      27.5%; Score 49; DB 2; Length 413;
Best Local Similarity 80.0%; Pred. No. 88;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 20 ALFWHPTYQL 29
Db 146 ALVWHPYQL 155

RESULT 9
US-09-543-681A-5864
; Sequence 5864, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5864
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5864

Query Match      27.5%; Score 49; DB 4; Length 446;
Best Local Similarity 47.8%; Pred. No. 96;
Matches 11; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 3 GTLWGSMEAWLATVLKALPWP 25
Db 321 GILKSGSLSDSLAVILSLNDMHP 343

RESULT 10
US-09-413-814-34
; Sequence 34, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloecker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hofle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or
; TITLE OF INVENTION: heteropolypeptide compounds
; FILE REFERENCE: PCT/US 99/23535
```

; CURRENT APPLICATION NUMBER: US/09/413,814
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-413-814-34

Query Match 27.5%; Score 49; DB 3; Length 557;

Best Local Similarity 47.6%; Pred. No. 1.2e+02;
Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 5 LAGSGMEAWLATVLKALPWP 25
Db 491 LHGLDPEAYLADVIRAMPWP 511

RESULT 11

US-09-489-039A-7417
; Sequence 7417, Application US/09489039A
; Patent No. 6610836

; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 7417

; LENGTH: 345

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-7417

Query Match 27.2%; Score 48.5; DB 4; Length 345;

Best Local Similarity 42.3%; Pred. No. 84;
Matches 11; Conservative 2; Mismatches 10; Indels 3; Gaps 1;

Qy 1 GQSTLWGSMEA---WLATVLKALPW 23
Db 29 GRAPLWGEGLAALWMLGGLFTLTW 54

RESULT 12

US-09-252-991A-30198
; Sequence 30198, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 30198

; LENGTH: 1239

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-30198

Query Match 27.2%; Score 48.5; DB 4; Length 1239;

Best Local Similarity 50.0%; Pred. No. 3.7e+02;
Matches 9; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

Qy 6 WSGSGMEAWLATVLKALPW 23
Db 15 W-SQIAWLVMVKPWPW 31

RESULT 13

US-09-248-796A-24595

; Sequence 24595, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 24595

; LENGTH: 208

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-24595

Query Match 27.0%; Score 48; DB 4; Length 208;

Best Local Similarity 44.4%; Pred. No. 54;
Matches 8; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 3 GTLWGSMEAWLATVLKA 20
Db 1 GFFWGGGFPWAGTALVA 18

RESULT 14

US-09-489-039A-13604

; Sequence 13604, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 13604

; LENGTH: 326

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-13604

Query Match 27.0%; Score 48; DB 4; Length 326;

Best Local Similarity 40.7%; Pred. No. 92;
Matches 11; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

Qy 4 TLMGSGMEAWLATVLKALPWHPTTYQL 30
Db 47 TLACGSSTAWRKDAVLAVLPQTLQOB 73

RESULT 15

US-09-252-991A-22806

; Sequence 22806, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

```

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22806
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22806

```

```

Query Match      26.7%; Score 47.5; DB 4; Length 80;
Best Local Similarity 57.9%; Pred. No. 21;
Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

```

```

QY      5 LWGSCMEA---WLATVLKA 20
      ||:||||| ||: ||
Db      15 LFGAGMEAAAPWLSLVLSA 33

```

Search completed: November 10, 2004, 14:55:42
Job time : 12.3972 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 35.8547 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-59

Perfect score: 178

Sequence: 1 GQGTLMGSGMEAWLATVLPKALPHPTQLEP 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap:
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pap:
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pap:
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pap:
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pap:
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pap:
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pap:
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap:
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pap:
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pap:
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pap:
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pap:
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pap:
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pap:
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pap:
16: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pap:
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pap:
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pap:
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pap:
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	178	100.0	31	14	US-10-092-750-59
2	62.5	35.1	241	16	US-10-437-963-133353
3	60.5	34.0	94	15	US-10-264-049-3360
4	60.5	34.0	526	16	US-10-437-963-164238
5	60	33.7	131	17	US-10-425-115-271630
6	55.5	31.2	569	15	US-10-424-599-279403
7	55.5	31.2	584	15	US-10-425-114-47346
8	55	30.9	853	16	US-10-437-963-124083
9	54.5	30.6	48	9	US-09-995-494-102
10	54.5	30.6	242	17	US-10-425-115-235453
11	54	30.3	74	17	US-10-425-115-228082
12	54	30.3	575	15	US-10-282-122A-62303
13	54	30.3	575	15	US-10-282-122A-64407

14	53.5	30.1	542	15	US-10-425-114-69836	Sequence 69836, A
15	53.5	30.1	607	17	US-10-425-115-198410	Sequence 198410, A
16	53.5	30.1	618	15	US-10-425-114-64802	Sequence 64802, A
17	53	29.8	11300	16	US-10-250-304A-2	Sequence 2, Appl
18	52.5	29.5	55	9	US-09-864-761-40496	Sequence 40496, A
19	52.5	29.5	63	9	US-09-864-761-41691	Sequence 41691, A
20	52	29.2	131	17	US-10-425-115-362054	Sequence 362054, A
21	52	29.2	218	15	US-10-425-114-53728	Sequence 53728, A
22	52	29.2	335	17	US-10-425-115-208290	Sequence 208290, A
23	52	29.2	350	15	US-10-425-114-58750	Sequence 58750, A
24	51.5	28.9	50	9	US-09-864-761-40675	Sequence 40675, A
25	51.5	28.9	224	14	US-10-104-047-2601	Sequence 2601, Ap
26	51.5	28.9	599	16	US-10-437-963-154605	Sequence 154605, A
27	51.5	28.9	763	13	US-10-081-408-2	Sequence 2, Appl
28	51.5	28.9	998	13	US-10-081-408-20	Sequence 20, Appl
29	51	28.7	130	15	US-10-424-599-266380	Sequence 266380, A
30	51	28.7	213	17	US-10-425-115-292104	Sequence 292104, A
31	51	28.7	214	15	US-10-282-122A-68380	Sequence 68380, A
32	51	28.7	312	16	US-10-437-963-132094	Sequence 132094, A
33	51	28.7	410	15	US-10-425-115-246019	Sequence 246019, A
34	51	28.7	455	17	US-10-425-115-246019	Sequence 246019, A
35	51	28.7	801	15	US-10-389-566-2305	Sequence 2305, Ap
36	51	28.7	863	17	US-10-425-115-307525	Sequence 307525, A
37	50.5	28.4	150	15	US-10-424-599-257786	Sequence 257786, A
38	50.5	28.4	567	16	US-10-437-963-183584	Sequence 183584, A
39	50.5	28.4	981	14	US-10-174-677-31	Sequence 31, Appl
40	50.5	28.4	981	15	US-10-380-705-2	Sequence 2, Appl
41	50.5	28.4	981	15	US-10-210-172-88	Sequence 88, Appl
42	50	28.1	74	15	US-10-424-599-218149	Sequence 218149, A
43	50	28.1	88	16	US-10-437-963-145388	Sequence 145388, A
44	50	28.1	111	15	US-10-424-599-145125	Sequence 145125, A
45	50	28.1	229	14	US-10-017-161-2148	Sequence 2148, Ap

ALIGNMENTS

RESULT 1
US-10-092-750-59
Sequence 59, Application US/10092750
Publication No. US20030032157A1
GENERAL INFORMATION:
APPLICANT: Hammond, Philip W.
APPLICANT: Alpin, Julia
APPLICANT: Wright, Martin C.
TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
FILE REFERENCE: 50036/050002
CURRENT APPLICATION NUMBER: US/10/092,750
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: US 60/274,526
PRIOR FILING DATE: 2001-03-08
NUMBER OF SEQ ID NOS: 253
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 59
LENGTH: 31
TYPE: PRT
ORGANISM: Homo sapiens
US-10-092-750-59

Query Match 100.0%; Score 178; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 6.3e-16;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GQGTLMGSGMEAWLATVLPKALPHPTQLEP 31
DB 1 GQGTLMGSGMEAWLATVLPKALPHPTQLEP 31
RESULT 2
US-10-437-963-133353
Sequence 133353, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:

```

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 135353
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; LOCATION: (1)..(241)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_37038C.1.pep
; US-10-437-963-135353

Query Match 35.1%; Score 62.5; DB 16; Length 241;
Best Local Similarity 47.1%; Pred. No. 5.1;
Matches 16; Conservative 4; Mismatches 11; Indels 3; Gaps 2;

QY 1 GQGTWLG-SGMEAWLATVVKALP--WHPTYQLEP 31
DB 32 GGGVWRCGSGGGLAGVRRSLPYLWPPTPQLPP 65

RESULT 3
US-10-264-049-3360
; Sequence 3360, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3360
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (2)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (76)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (83)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (93)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; US-10-264-049-3360

Query Match 34.0%; Score 60.5; DB 15; Length 94;
Best Local Similarity 45.8%; Pred. No. 3.6;
Matches 11; Conservative 1; Mismatches 7; Indels 5; Gaps 1;

QY 1 GQGTWLGSGMEAWLATVVKALPWH 24
DB 38 GRNFWGCGWEAWC-----LLRWH 56

RESULT 4
US-10-437-963-164238
; Sequence 164238, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 164238
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_63157C.1.pep
; US-10-437-963-164238

Query Match 34.0%; Score 60.5; DB 16; Length 526;
Best Local Similarity 38.1%; Pred. No. 20;
Matches 16; Conservative 5; Mismatches 6; Indels 15; Gaps 3;

QY 1 GQGTWLGSGM-----EA--WLATVVKALPWH-----PTY 27
DB 252 GEGSQSGRLLLPPEANMTALVIEASPMWRRTDVAVPYTY 293

RESULT 5
US-10-425-115-271630
; Sequence 271630, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 271630
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_17931C.1.pep
; US-10-425-115-271630

Query Match 33.7%; Score 60; DB 17; Length 131;
Best Local Similarity 38.7%; Pred. No. 5.8;
Matches 12; Conservative 5; Mismatches 12; Indels 2; Gaps 1;

QY 1 GQGTWLGSGMEAWLATVVKALPWHPTYQLEP 31
```

Db 22 GGRRWAVG--SFVVTINGATPWLPPRSAP 50

RESULT 6

US-10-424-599-279403
; Sequence 279403, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 279403

; LENGTH: 569
; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_94322C.1.pep
US-10-424-599-279403

Query Match 31.2%; Score 55.5; DB 15; Length 569;

Best Local Similarity 52.0%; Pred. No. 98;

Matches 13; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY 2 GQTLWGSQMEAWL-ATVLKALPWHPP 25

Db 121 EGTLLWHAHSSWLRATVYGALIHP 145

RESULT 7

US-10-425-114-47346

; Sequence 47346, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 47346

; LENGTH: 584
; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: 701043069_FLI.pep
US-10-425-114-47346

Query Match

Best Local Similarity 31.2%; Score 55.5; DB 15; Length 584;

Matches 13; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY 2 GQTLWGSQMEAWL-ATVLKALPWHPP 25

Db 136 EGTLLWHAHSSWLRATVYGALIHP 160

RESULT 8

US-10-437-963-124083

; Sequence 124083, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 124083

; LENGTH: 853

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_26857C.1.pep

US-10-437-963-124083

Query Match 30.9%; Score 55; DB 16; Length 853;

Best Local Similarity 37.0%; Pred. No. 1.7e+02;

Matches 10; Conservative 2; Mismatches 13; Indels 2; Gaps 1;

QY 1 GQTLWGSQMEAWL-ATVLKALPWHPP 27

Db 222 GNVTIWRSQ--WTGQNFVGIWPRPLY 246

RESULT 9

US-09-995-494-102

; Sequence 102, Application US/09995494

; Patent No. US20020127578A1

; GENERAL INFORMATION:

; APPLICANT: Salceda, Susana

; APPLICANT: Macina, Roberto

; APPLICANT: Recipon, Herve

; APPLICANT: Cafferkey, Robert

; APPLICANT: Ali, Shujath

; APPLICANT: Sun, Yongming

; APPLICANT: Liu, Chenghua

; APPLICANT: Chen, Sei-Yu

; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and

; FILE REFERENCE: DEX-0293

; CURRENT APPLICATION NUMBER: US/09/995,494

; CURRENT FILING DATE: 2001-11-27

; PRIOR APPLICATION NUMBER: 60/253,176

; NUMBER OF SEQ ID NOS: 115

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 102

; LENGTH: 48

; TYPE: PRT

; ORGANISM: Homo sapien

; OTHER INFORMATION: 102

US-09-995-494-102

Query Match

Best Local Similarity 30.6%; Score 54.5; DB 9; Length 48;

Matches 13; Conservative 3; Mismatches 9; Indels 5; Gaps 2;

QY 1 GQTLW-GSGMEAWL-ATVLKALPWHPPYQL 29

Db 17 GIGLWWSGRGTN---TPVKSIGWHKTYQL 42

RESULT 10

US-10-425-115-235453

; Sequence 235453, Application US/10425115

; Publication No. US20040214272A1

```
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 235453
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_146314C.1.pep
US-10-425-115-235453

Query Match          30.6%; Score 54.5; DB 17; Length 242;
Best Local Similarity 40.0%; Pred. No. 56;
Matches 12; Conservative 4; Mismatches 13; Indels 1; Gaps 1;

QY 1 GGGTGWGCGMEAWLATVLKALPWHPTTYOLE 30
      |||||
Db 183 GSGKVAWGTAACATYLECKYHPTMEAE 211

RESULT 11
US-10-425-115-228082
; Sequence 228082, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 228082
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_139605C.1.pep
US-10-425-115-228082

Query Match          30.3%; Score 54; DB 17; Length 74;
Best Local Similarity 46.4%; Pred. No. 19;
Matches 13; Conservative 4; Mismatches 5; Indels 5; Gaps 2;

QY 6 WSGGMEAWLATVLKALPWHPTTYQL 29
      |||||
Db 23 WSGGVNGSTAGWPASLIKALP--PPAQL 48

RESULT 12
US-10-282-122A-62303
; Sequence 62303, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
```

```
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 62303
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Mycobacterium bovis
US-10-282-122A-62303

Query Match          30.3%; Score 54; DB 15; Length 575;
Best Local Similarity 38.5%; Pred. No. 1.6e+02;
Matches 10; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

QY 6 WSGGMEAWLATVLKALPWHPTTYQLEP 31
      |||||
Db 422 WNRVQAWLABETGQPPWTFWAGRP 447

RESULT 13
US-10-282-122A-64407
; Sequence 64407, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
```

```
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64407
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64407

Query Match      30.3%; Score 54; DB 15; Length 575;
Best Local Similarity 38.5%; Pred. No. 1.6e+02;
Matches 10; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

QY      6 WSGGMEAWLATVLKALPWHPTYQLEP 31
Db      422 WNRVQWLAEETGQPPWTPWYAGRP 447

RESULT 14
US-10-425-114-69836
; Sequence 69836, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 69836
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROB73006B02_FIL.pep
US-10-425-114-69836

Query Match      30.1%; Score 53.5; DB 15; Length 542;
Best Local Similarity 52.0%; Pred. No. 1.7e+02;
Matches 13; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY      2 OGTLWGSMEAWL-ATVLKALPWH 25
Db      54 EGTLWWHAHSSWLRAVHGALIHP 88

RESULT 15
US-10-425-115-198410
; Sequence 198410, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
```

```
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 198410
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(607)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_112527C.1.pep
US-10-425-115-198410

Query Match      30.1%; Score 53.5; DB 17; Length 607;
Best Local Similarity 52.0%; Pred. No. 1.9e+02;
Matches 13; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY      2 OGTLWGSMEAWL-ATVLKALPWH 25
Db      129 EGTLWWHAHSSWLRAVHGALIHP 153

Search completed: November 11, 2004, 02:43:15
Job time : 36.9047 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run On: November 10, 2004, 13:40:53 ; Search time 7.60377 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-59
Perfect score: 178

Sequence: 1 GQGTLWGSMEAWLATVTKALPWHPTTYQLEP 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59.5	33.4	479	2 S68598	sucrose-6-phosphat
2	56.5	31.7	477	2 A33647	melittin resistanc
3	56	31.5	304	2 A10722	conserved hypothet
4	56	31.5	614	2 T09902	hypothetical prote
5	55	30.9	590	2 A86307	amino acid transpo
6	54	30.3	575	2 B70612	probable recD prot
7	54	30.3	576	2 S12792	protein-tyrosine k
8	54	30.3	888	2 I58378	tyrosine kinase -
9	53.5	30.1	454	2 A43501	sucrose-6-phosphat
10	53.5	30.1	762	2 A54411	amine oxidase (cop
11	51.5	28.9	559	2 T48180	hypothetical prote
12	51.5	28.9	763	2 JC5234	amine oxidase (cop
13	51	28.7	25	2 P70328	Ig heavy chain CDR
14	51	28.7	848	1 T02053	Receptor kinase
15	50	28.1	402	2 A18553	sodium-dependent n
16	50	28.1	464	2 A43625	protein-tyrosine k
17	50	28.1	480	2 T20899	hypothetical prote
18	50	28.1	577	2 T45333	hypothetical prote
19	50	28.1	753	2 G87178	probable DNA-bind
20	50	28.1	864	2 A48266	protein-tyrosine k
21	49.5	27.8	286	2 B72404	Glycine-CRNA ligas
22	49.5	27.8	386	2 T14243	ubiquinol-cytochro
23	49	27.5	611	1 S06047	endo-1,4-beta-xyla
24	48.5	27.2	349	2 A52369	hypothetical prote
25	48.5	27.2	385	2 S53087	ubiquinol-cytochro
26	48.5	27.2	749	2 S71715	sensory transducti
27	48	27.0	88	2 I40653	9K protein-Coxie
28	48	27.0	179	2 A03484	phosphinothricin N
29	48	27.0	227	2 D84366	mannose-1-phosphat

30 48 27.0 248 2 T00300
31 48 27.0 248 2 JQ1340
32 48 27.0 294 2 AC3901
33 48 27.0 304 2 G08339
34 48 27.0 304 2 G85697
35 48 27.0 344 2 E64865
36 48 27.0 344 2 AH0041
37 48 27.0 374 2 C84339
38 48 27.0 475 2 AC1246
39 48 27.0 475 2 AG1608
40 48 27.0 503 2 E87101
41 48 27.0 614 2 S64447
42 48 27.0 622 2 S61140
43 48 27.0 831 2 S76235
44 47.5 26.7 142 2 T51902
45 47.5 26.7 251 1 ESDKTM

trax protein - Esc
trax protein - Esc
conserved hypothet
L, D-carboxyprptid
L, D-carboxyprptid
Muramoyltetrapepti
L-rhamnose-proton
quinolinate synthe
branched-chain alp
probable membrane
probable membrane
hypothetical prote
hypothetical prote
oleoyl-[acyl-carri

ALIGNMENTS

RESULT 1

S68598 sucrose-6-phosphate hydrolase ScrB - Streptococcus sobrinus (strain 6715)
C;Species: Streptococcus sobrinus

A;Variety: Strain 6715
C;Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 07-May-1999

C;Accession: S68598
R;Chen, Y.Y.M.; Lee, L.N.; LeBlanc, D.J.

Infect. Immun. 61, 2602-2610, 1993
A;Title: Sequence analysis of scrA and scrB from Streptococcus sobrinus 6715.

A;Reference number: S68598; MUID:93273516; PMID:8500898
A;Accession: S68598

A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-479 <CHE>
A;Cross-references: EMBL:L06792

C;Genetics:
A;Gene: scrB

C;Superfamily: sucrose-6-phosphate hydrolase

Query Match 33.4%; Score 59.5; DB 2; Length 479;
Best Local Similarity 39.3%; Pred. No. 4.3;
Matches 11; Conservative 3; Mismatches 7; Indels 7; Gaps 1;

Qy 11 EAWLATVTKAL-----PWHPTTYQLEP 31
Db 13 EDWSAEIKTIQENWAKSPWHPTTYHIEP 40

RESULT 2

AB3647 melittin resistance protein pqxAB [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis

C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AB3647

R;DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Lete

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melite

A;Reference number: AD3252; PMID:11756688
A;Accession: AB3647

A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-477 <KUR>
A;Cross-references: UNIPROT:Q8YB02; GB:AE008918; PIDN:AAL54341.1; PID:gl7985322; GSPDB:

A;Experimental source: strain 16M.
C;Genetics:

A;Gene: BMEI11099
A;Map position: II

Query Match 31.7%; Score 56.5; DB 2; Length 477;
Best Local Similarity 39.3%; Pred. No. 11;

Matches 11; Conservative 4; Mismatches 8; Indels 5; Gaps 1;
 QY 3 GTLWGSME-----AWLATVVKALPWP 25
 Db 248 GDLYGAGREHARGSIWLFVATLWSP 275

RESULT 3
 AIO722
 conserved hypothetical protein STY1928 [imported] - Salmonella enterica subsp. enterica
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: This species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AIO722
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Mouton, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Park, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AIO722
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-304 <PAR>
 A:Cross-references: GB:ALJ513382; PIDN:CAD05483.1; PID:GL6502987; GSPDB:GN00176
 A:Genetics:
 C:Gene: STY1928

Query Match 31.5%; Score 56; DB 2; Length 304;
 Best Local Similarity 40.0%; Pred. No. 7.8;
 Matches 10; Conservative 4; Mismatches 9; Indels 2; Gaps 1;
 QY 2 QGTLWGSMEAWLATVVKALPWP 26
 Db 169 QGTLWGSNLAFLISLI-GTPWNP 191

RESULT 4
 T09902
 hypothetical protein T22A6.220 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
 C:Accession: T09902
 R:Bevan, M.; Zimmermann, W.; Gruenewald, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; May
 submitted to the Protein Sequence Database, June 1999
 A:Reference number: Z16896
 A:Accession: T09902
 A:Molecule type: DNA
 A:Residues: 1-614 <BEV>
 A:Cross-references: UNIPROT:O9STV5; EMBL:AL078637; GSPDB:GN00062; ATSP:T22A6.220
 A:Experimental source: cultivar Columbia; BAC clone T22A6
 C:Genetics:
 A:Gene: ATSP:T22A6.220
 A:Map position: 4
 A:Introns: 192/2; 358/3

Query Match 31.5%; Score 56; DB 2; Length 614;
 Best Local Similarity 44.4%; Pred. No. 16;
 Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 6 WGSMEAWLATVVKALPWP 23
 Db 125 WGAQFSPWVAATAKAYPW 142

RESULT 5
 A86307
 amino acid transporter homolog [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: A86307
 R:theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Marziali
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Venter, J.C.; Davis, R.W.
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: A86307
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-590 <STO>
 A:Cross-references: UNIPROT:Q9SHH0; GB:AE005172; NID:G5734765; PIDN:AAD50030.1; GSPDB:G
 C:Genetics:
 A:Map position: 1
 C:Superfamily: ecotropic retrovirus receptor protein

Query Match 30.9%; Score 55; DB 2; Length 590;
 Best Local Similarity 34.8%; Pred. No. 21;
 Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
 QY 1 GQGTLMGSGMEAWLATVVKALPW 23
 Db 482 GVSALWNSGVKGMIAVTVTGVW 504

RESULT 6
 B70612
 probable recD protein - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C:Accession: B70612
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: B70612
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-575 <COL>
 A:Cross-references: UNIPROT:P96919; GB:292772; GB:AL123456; NID:G3261722; PIDN:CAB07118
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: recD
 C:Superfamily: exodeoxyribonuclease V 67K chain

Query Match 30.3%; Score 54; DB 2; Length 575;
 Best Local Similarity 38.5%; Pred. No. 27;
 Matches 10; Conservative 2; Mismatches 14; Indels 0; Gaps 0;
 QY 6 WGSMEAWLATVVKALPWP 31
 Db 422 WNRVQAWLAETGQPPWTPWYGRP 447

RESULT 7
 S12792
 protein-tyrosine kinase (EC 2.7.1.112) ltk - mouse
 N:Alternate names: leukocyte tyrosine kinase
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-2004
 C:Accession: S12792; S00904
 R:Bernards, A.; de la Monte, S.M.
 EMBO J. 9, 2279-2287, 1990
 A:Title: The ltk receptor tyrosine kinase is expressed in pre-B lymphocytes and cerebra
 A:Reference number: S12792; MUID:90291994; PMID:2357970
 A:Accession: S12792
 A:Molecule type: mRNA

R.; Sato, Y.; Poy, F.; Jacobson, G.R.; Kuramitsu, H.K.
J. Bacteriol. 171, 263-271, 1989
#Title: Characterization and sequence analysis of the scrA gene encoding enzyme II (scrA)
#Accession: A32243 #MIM:253656
#Reference number: A32243 #UID:89123027; PMID:253656

Query Match 30.1%; Score 53.5; DB 2; Length 454;
Best Local Similarity 34.6%; Pred. No. 25;
Matches 9; Conservative 4; Mismatches 6; Indels 7; Gaps 1;

QY 13 WLATVLKAL-----PWHTYOLEP 31
 | : : : ||| ||| : : :
DDB 15 WTBEIKSITKNVALSPWHTTHIEP 40

RESULT 10

A54411
amine oxidase (copper-containing) (EC 1.4.3.6), serum, precursor - bovine
C-species: Bos primigenius taurus (cattle)
Date: 09-Sep-1994 #sequence revision 06-Feb-1995 #text_change 09-Jul-2004
A54411; #Accession: A54411; #Release: A8242; S65408
R.; Mu, D.; Medhinradsky, K.F.; Adams, G.W.; Mayer, P.; Hines, W.M.; Burlingame, A.L.; S
J. Biol. Chem. 269, 9926-9932, 1994
A54411; Title: Primary structures for a mammalian cellular and serum copper amine oxidase.
A-reference number: A54411; UID:94193686; PMID:8144587
A-accession: A54411
A-molecule type: mRNA
A-residues: 1-762 <MUA>
R.; Mu, D.; Janes, S.M.; Smith, A.J.; Brown, D.E.; Docley, D.M.; Klinman, J.P.
J. Biol. Chem. 267, 7979-7982, 1992
A54411; Tyrosine codon corresponds to topa quinone at the active site of copper amine
A-reference number: A38081; UID:92235001; PMID:1569055
A-accession: B38081
A-molecule type: protein
A-residues: 463-465,'D','I',467-473,'X',475-485 <MU2>
R.; Janes, S.M.; Mu, D.; Wenner, D.; Smith, A.J.; Kaur, S.; Maltby, D.; Burlingame, A.L.; S
Science 248, 981-987, 1990
A54411; A new redox cofactor in eukaryotic enzymes: 6-hydroxydopa at the active site c
A-reference number: A48242; UID:90260648; PMID:2111581
A-accession: A48242
A-status: preliminary
A-molecule type: protein
A-residues: 468-472 <JAN>
Ride Blase, D.; Acostinelli, E.; de Matteis, G.; Mondovi, B.; Morpurgo, L.
Eur. J. Biochem. 237, 93-99, 1996
A54411; Half-of-the-sites reactivity of bovine serum amine oxidase. Reactivity and ch
A-reference number: S65408; UID:96203913; PMID:8620899
A-accession: S65408
A-molecule type: protein
A-residues: 463-469,'X',471-487 <DEB>
C-suprafamily: amiloride-binding protein
C-keywords: copper; glycoprotein; oxidoreductase; quinoicoprotein; topaquinoxene
F17-16/Domains: signal sequence #status predicted <SIG>
F17-762/Product: amine oxidase (copper-containing) #status predicted <MAT>
F136,231,665/binding site: carboxylate (Asn) (covalent) #status predicted
F1470/Modified site: topaquinoxene (Tyr) #status experimental

Query Match 30.1%; Score 53.5; DB 2; Length 762;
Best Local Similarity 37.0%; Pred. No. 42;
Matches 10; Conservative 6; Mismatches 6; Indels 5; Gaps 1;

QY 9 GNEAW-----LATVTKALPMHPTYOLE 30
 | : : : ||| ||| : : :
Db 533 GLENVVWAEDMAFVTAPWSPEHQIQ 559

RESULT 11

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: A11853
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-402 <KUR>
A;Cross-references: UNIPROT:Q8YZS9; GB:BA000019; PIDN:BA072336.1; PID:g17129723; GSPDB:C
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all0378
C;Superfamily: pyrimidine nucleoside transport protein nupC

Query Match	28.1%	Score 50;	DB 2;	Length 402;
Best Local Similarity	50.0%;	Pred. No. 62;		
Matches	9;	Conservative 2;	Mismatches 7;	Indels 0;
			Gaps	0;

Oy 6 WSGGMEALATVILKALPW 23
|||:|:|:|:
Db 34 WGLGLEFVALVILKTPW 51

Search completed: November 10, 2004, 14:52:26
Job time : 7.60377 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 : Search time 41.2943 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-59

Perfect score: 178

Sequence: 1 GQTLWNGMEAWLATVLPWHPYQLP 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_spriot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60.5	34.0	565	2 Q9FLB5	Q9flb5 arabidopsis
2	57	32.0	199	2 Q8RVJ4	Q8rvj4 pinus pinas
3	56.5	31.7	477	2 Q8YB02	Q8yb02 bruceella me
4	56.5	31.7	478	2 Q8FXE2	Q8fxe2 bruceella su
5	56	31.5	304	2 Q8Z689	Q8z689 salmonella
6	56	31.5	304	2 Q8ZP18	Q8zp18 salmonella
7	56	31.5	614	2 Q9STV5	Q9stv5 arabidopsis
8	56	31.5	623	2 Q8RWQ8	Q8rwq8 arabidopsis
9	55	30.9	305	2 Q6SL32	Q6sl32 armillifer
10	55	30.9	305	2 AAS00841	Aas00841 armillifer
11	55	30.9	590	2 Q9SHH0	Q9shh0 arabidopsis
12	55	30.9	599	2 Q7WMJ5	Q7wmj5 porphyromon
13	55	30.9	853	2 Q7EZ32	Q7ez32 oryza sativ
14	54.5	30.6	952	2 Q6Q0N0	Q6q0n0 rattus norv
15	54.5	30.6	952	2 AAS75317	Aas75317 rattus no
16	54	30.3	307	2 Q7SC19	Q7sc19 neurospora
17	54	30.3	561	2 Q7D914	Q7d914 mycobacteri
18	54	30.3	575	2 P96919	P96919 mycobacteri
19	54	30.3	575	2 Q7U1L1	Q7u1l1 mycobacteri
20	54	30.3	707	2 Q98TY6	Q98ty6 rhizobium 1
21	54	30.3	888	1 KLTK_MOUSE	P08923 mus musculu
22	53.5	30.1	418	2 Q7JCD2	Q7jcd2 bacillus ce
23	53.5	30.1	418	2 AAS40064	Aas40064 bacillus
24	53.5	30.1	479	1 SCRB_STRMU	P13522 streptococc
25	53.5	30.1	762	1 AOCK_BOVIN	Q29437 bos taurus
26	53	29.8	341	2 Q7UWM6	Q7uwm6 rhodospirell
27	53	29.8	412	2 Q8VED1	Q8ved1 mus musculu
28	53	29.8	439	2 Q8FFJ4	Q8ffj4 arabidopsis
29	52.5	29.5	240	2 Q9DZ20	Q9dzz0 mus musculu
30	52.5	29.5	553	2 Q87IG3	Q87ig3 vibrio para
31	52.5	29.5	734	1 PSAB_CYACA	Q9tlq6 cyanidium c

32 52.5 29.5 944 1 VGLB_TUHV2
33 52.5 29.5 944 2 Q77L59
34 52.5 29.5 944 2 AAK57104
35 52 29.2 166 2 Q6XCA0
36 52 29.2 166 2 AAP04340
37 52 29.2 227 2 Q8HMB2
38 52 29.2 363 2 Q9QTE4
39 52 29.2 437 2 Q9E942
40 52 29.2 468 2 Q7UN7
41 52 29.2 471 2 Q7UR92
42 52 29.2 883 2 Q9RF11
43 51.5 28.9 238 2 Q76E57
44 51.5 28.9 238 2 BAD12557
45 51.5 28.9 240 2 Q804J1

Q9wrl5 tupaiid herp
Q77l59 tupaiid herp
Aak57104 tupaiid he
Q6xca0 mustelus ma
Aap04340 mustelus
Q8hmb2 lamprogramm
Q9qte4 marek's dis
Q9e942 gallid herp
Q7un7 synchococc
Q7ur92 rhodopirell
Q9rf11 myxococcus
Q76e57 homo sapien
Bad12557 homo sapi
Q804j1 brachydanio

ALIGNMENTS

RESULT 1

Q9FLB5 PRELIMINARY; PRT; 565 AA.
AC Q9FLB5
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Laccase (Diphenol oxidase).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=98344145; PubMed=9679202;
RX Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
RA Tabata S.;
RA "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
features of the regions of 1,381,565 bp covered by twenty one
physically assigned P1 and TAC clones."
RL DNA Res. 5:131-145(1998).
DR EXBL; AB010692; BAB09982.1; -.
DR HSSP; Q9Y780; IHFU.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR002355; Cu ox copper_BS.
DR Pfam; PF00394; Cu-oxidase; 1.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 1.
DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
SQ SEQUENCE 565 AA; 62733 MW; 15DA9DD5CA8932E CRC64;

Query Match 34.0%; Score 60.5; DB 2; Length 565;
Best Local Similarity 53.8%; Pred. No. 23;
Matches 14; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

Qy 2 QGTLWNGMEAWLATVLPWHPYQLP 26

Db 121 EGTLMWHAHSHSLRATVYGALIIHPT 146

RESULT 2

Q8RVJ4 PRELIMINARY; PRT; 199 AA.
AC Q8RVJ4
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Putative acid phosphatase (Fragment).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=71647;

```

RN RP SEQUENCE FROM N.A.
RC TISSUE=Aerial part;
RA Dubos C., Le Provost G., Salin F., Lalane C., Madur D., Frigerio J.M.,
RL Plomion C.;
RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL; AJ309082; CAC8485.1; -.
DR InterPro; IPR005519; acid_phosphat_B.
DR Pfam; PF03767; Acid_phosphat_B; 1.
FT NON_TER 1
FT 199
SQ SEQUENCE 199 AA; 22649 MW; 082799D90C773DD5 CRC64;

Query Match 32.0%; Score 57; DB 2; Length 199;
Best Local Similarity 37.9%; Pred. No. 24;
Matches 11; Conservative 7; Mismatches 7; Indels 4; Gaps 1;

QY 4 TLWGSGMEAWL---ATVLKALPWPHTYQ 28
DB 41 TLAGDGDARVDPADETLLSNIPYENYE 69

RESULT 3
Q8YB02 PRELIMINARY; PRT; 477 AA.
AC Q8YB02;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DE MELITTIN RESISTANCE PROTEIN PQAB.
GN OrderedLocusNames=EMBI11099;
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;

SEQUENCE FROM N.A.
RP STRAIN=16M / ATCC 23456 / Biotype 1;
RC MEDLINE=20020109; PubMed=11756688; DOI=10.1073/pnas.221575398;
RA Delvecchio V.G., Kaputal V., Redkar R.J., Patra G., Mijer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA Seikov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyriades N.C., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR ENBL; AB009741; AAL54341.1; -.
DR PIR; AB3647; AB3647.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0000030; F:mannosyltransferase activity; IEA.
DR GO; GO:0006493; P:O-linked glycosylation; IEA.
DR InterPro; IPR003342; Glyco_trans_39.
DR Pfam; PF02366; PMT; 1.
KW Complete proteome;
SQ SEQUENCE 477 AA; 51822 MW; 1DSF2CB34344D96E CRC64;

Query Match 31.7%; Score 56.5; DB 2; Length 477;
Best Local Similarity 39.3%; Pred. No. 65;
Matches 11; Conservative 4; Mismatches 8; Indels 5; Gaps 1;

QY 3 GTLWGSME-----AWLATVLKALPWP 25
DB 248 GDLYGAGREHARGSIWLFWIVATLPWSP 275

RESULT 4
Q8FXE2 PRELIMINARY; PRT; 478 AA.
AC Q8FXE2;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)

SEQUENCE FROM N.A.
RP STRAIN=16M / ATCC 23456 / Biotype 1;
RC MEDLINE=20020109; PubMed=11756688; DOI=10.1073/pnas.221575398;
RA Delvecchio V.G., Kaputal V., Redkar R.J., Patra G., Mijer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA Seikov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyriades N.C., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR ENBL; AB009741; AAL54341.1; -.
DR PIR; AB3647; AB3647.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0000030; F:mannosyltransferase activity; IEA.
DR GO; GO:0006493; P:O-linked glycosylation; IEA.
DR InterPro; IPR003342; Glyco_trans_39.
DR Pfam; PF02366; PMT; 1.
KW Complete proteome;
SQ SEQUENCE 477 AA; 51822 MW; 1DSF2CB34344D96E CRC64;

Query Match 31.7%; Score 56.5; DB 2; Length 477;
Best Local Similarity 39.3%; Pred. No. 65;
Matches 11; Conservative 4; Mismatches 8; Indels 5; Gaps 1;

QY 3 GTLWGSME-----AWLATVLKALPWP 25
DB 248 GDLYGAGREHARGSIWLFWIVATLPWSP 275

RESULT 5
Q8Z689 PRELIMINARY; PRT; 304 AA.
AC Q8Z689; Q7CAF6;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE Hypothetical protein STY1928.
GN OrderedLocusNames=STY1928, t1077;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
EN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=CT18;
RC MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Kragh A., Larsen T.S., Leather S., Moulé S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Ty2 / ATCC 700931;
RC MEDLINE=22531367; PubMed=12644504;

```

DE Dolichyl-phosphate-mannose-protein mannosyltransferase family protein.

DE OrderedLocusNames=BRA0137;

OS Brucella suis.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Brucellaceae; Brucella.

OX NCBI_TaxID=29461;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1330 / Biovar 1;

RC MEDLINE=22247741; PubMed=12271122; DOI=10.1073/pnas.192319099;

RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,

RA Read T.D., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,

RA Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R.,

RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J.A.,

RA Van Aken S.E., Riedmuller S., Tettelin H., Gill S.R., White O.,

RA Salzberg S.L., Hoover D.L., Lindler L.E., Hailing S.M., Boyle S.M.,

RA Fraser C.M.;

RT "The Brucella suis genome reveals fundamental similarities between

RT animal and plant pathogens and symbionts.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).

DR ENBL; AE014517; AAN3345.1; -.

DR TIGR; BRA0137; -.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0000030; F:mannosyltransferase activity; IEA.

DR GO; GO:0016757; F:transferase activity, transferring glycosyl. .; IEA.

DR GO; GO:0006493; P:O-linked glycosylation; IEA.

DR InterPro; IPR003342; Glyco_trans_39.

DR Pfam; PF02366; PMT; 1.

KW Complete proteome; Glycosyltransferase; Transferase.

SQ SEQUENCE 478 AA; 51957 MW; AB668D96AFCBCE1 CRC64;

Query Match 31.7%; Score 56.5; DB 2; Length 478;

Best Local Similarity 39.3%; Pred. No. 65;

Matches 11; Conservative 4; Mismatches 8; Indels 5; Gaps 1;

QY 3 GTLWGSME-----AWLATVLKALPWP 25

DB 249 GDLYGAGREHARGSIWLFWIVATLPWSP 276

```

RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
DR EMBL; AL627272; CAD05483.1; -.
DR MEROPS; U61.001; -.
DR InterPro; IPR003507; Peptidase U61.
DR Pfam; PF02016; Peptidase U61; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 304 AA; 33365 MW; 26D9EA86D2DEC88E CRC64;

Query Match 31.5%; Score 56; DB 2; Length 304;
Best Local Similarity 40.0%; Pred. No. 49;
Matches 10; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

Qy 2 QGTLWGSMEALATLVKALPWHPT 26
Db 169 QGTLWGSMEALATLVKALPWHPT 191

RESULT 6
Q8ZP18 PRELIMINARY; PRT; 304 AA.
AC Q8ZP18;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative resistance protein MccP.
GN Name=ygqQ; OrderedLocusNames=STM1800;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen K., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea R.I., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
DR EMBL; AE008780; AAL20715.1; -.
DR MEROPS; U61.001; -.
DR InterPro; IPR003507; Peptidase U61.
DR Pfam; PF02016; Peptidase U61; 1.
KW Complete proteome.
SQ SEQUENCE 304 AA; 33351 MW; 2CEB67DDC06DD6D5 CRC64;

Query Match 31.5%; Score 56; DB 2; Length 304;
Best Local Similarity 40.0%; Pred. No. 49;
Matches 10; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

Qy 2 QGTLWGSMEALATLVKALPWHPT 26
Db 169 QGTLWGSMEALATLVKALPWHPT 191

RESULT 7
Q9STV5 PRELIMINARY; PRT; 614 AA.
AC Q9STV5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Transport inhibitor response-like protein.
GN Name=T22A6.220; Synonyms=AT4g24390;
OS Arabidopsis thaliana (Mouse-ear cross).

```

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Zimmermann W., Gruenisen A., Wambutt R., Bancroft I.,
RA Mewes H.W., Mayer K.F.X., Lemcke K., Schussler C.;
RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Zimmermann W., Gruenisen A., Wambutt R., Kalicki J., Wohldmann P.,
RA Smith A., Mewes H.W., Lemcke K., Mayer K.F.X.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; ALQ78637; CAB45074.1; -.
DR EMBL; AL161561; CAB79349.1; -.
DR PIR; T09902; T09902.
DR InterPro; IPR001810; P-box.
DR Pfam; PF00646; P-box; 1.
SQ SEQUENCE 614 AA; 68520 MW; A856CC01DF54ADB9 CRC64;

Query Match 31.5%; Score 56; DB 2; Length 614;
Best Local Similarity 44.4%; Pred. No. 96;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 6 WGSQMEALATLVKALPW 23
Db 125 WGAQFSPVAATAKAYEW 142

RESULT 8
Q8RWQ8 PRELIMINARY; PRT; 623 AA.
AC Q8RWQ8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE AT4g24390/T22A6.220.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Dale J.M., Hayashizaki Y., Heuan V.W., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RA Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY091772; AAM10320.1; -.
DR EMBL; BT006340; AAP21148.1; -.

```

[illegible]

Query Match 30.9%; Score 55; DB 2; Length 590;
Best Local Similarity 34.8%; Pred. No. 1.2e+02;
Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 GQTLWGSMEAWLATVLKALPW 23
DB 482 GVSALWNSGVKGIATVTGVIV 504

RESULT 12

Q7MWJ5 PRELIMINARY; PRT; 599 AA.
AC Q7MWJ5
DT 01-VAR-2004 (TREMBLrel. 26, Created)
DT 01-VAR-2004 (TREMBLrel. 26, Last sequence update)
DT 01-VAR-2004 (TREMBLrel. 26, Last annotation update)
DE GTP-binding protein TypA.
GN Name=typA; OrderedLocNames=P00615;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (Class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W83;
RX MEDLINE=22829867; PubMed=12949112;
RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
Dewhirst F.E., Fraser C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
Porphyromonas gingivalis strain W83.";
RL J. Bacteriol. 183:5591-5601(2003).
DR EMBL; AE017174; AAQ65799.1; --
DR TIGR; F06615; --
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0003746; P:translation elongation factor activity; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR GO; GO:0006414; P:translational elongation; IEA.
DR InterPro; IPR000640; EFG C
DR InterPro; IPR004461; EFTU_D2.
DR InterPro; IPR000795; ProtSyn_GTPbind.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR006298; trPA.
DR Pfam; PF00679; EFG C; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PRO0315; ELONGATNCT.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR TIGRFAMs; TIGR01394; typA_BipA; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW Complete proteome; GTP-binding.
SQ SEQUENCE 599 AA; 67195 MW; A3FC5E93046824E3 CRC64;

Query Match 30.9%; Score 55; DB 2; Length 599;
Best Local Similarity 35.7%; Pred. No. 1.3e+02;
Matches 10; Conservative 6; Mismatches 8; Indels 4; Gaps 1;

QY 4 TLWGSMEAWLATVLKALPWHPTQLEP 31
DB 162 TIYSAKQGWSTDYK---HPTQDIAP 185

RESULT 13

Q7EZ32 PRELIMINARY; PRT; 853 AA.
AC Q7EZ32
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Putative S-receptor kinase XKI1.

GN Name=OSUNB0041B22.1.09;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AP005126; BAC65367.1; --
DR GO; GO:0016301; F:Kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR001480; E lectin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR InterPro; IPR008271; Ser_Thr_kin_AS.
DR InterPro; IPR000858; Slocus_glycop.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF01453; B_lectin; 1.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00954; S locus glycop; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00108; B_lectin; 1.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00927; BULB_LLECTIN; 1.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS50948; PAN; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Receptor; Serine/threonine-protein kinase;
KW Transferrase.
SQ SEQUENCE 853 AA; 94610 MW; 2294AD9F81CF29B CRC64;

Query Match 30.9%; Score 55; DB 2; Length 853;
Best Local Similarity 37.0%; Pred. No. 1.8e+02;
Matches 10; Conservative 2; Mismatches 13; Indels 2; Gaps 1;

QY 1 GQTLWGSMEAWLATVLKALPWHPT 27
DB 222 GNVTVRSQ--WTGNFVGIPWRPLY 246

RESULT 14

Q6Q0N0 PRELIMINARY; PRT; 952 AA.
AC Q6Q0N0
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Calsynthenin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar;
RA Ahmed F., Torrado M., Zinovleva R.D., Senatorov V., Wistow G.,
Tomarev S.I.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 2 cadherin domains.
DR EMBL; AY569014; AAS75317.1; --
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR008985; ConA_like_lec_gl.
DR Pfam; PF00028; Cadherin; 2.
DR PRINTS; PR00205; CADHERIN.

Fri Nov 12 14:55:46 2004

us-10-092-750-59.rup

```
DR SMART; SW00112; CA; 2.  
DR PROSITE; PSS0288; CADHERIN_2; 2.  
KW Calcium; Cation-binding.  
SQ SEQUENCE 952 AA; 106260 MW; DBE2185AF20F9713 CRC64;  
  
Query Match      30.6%; Score 54.5; DB 2; Length 952;  
Best Local Similarity 50.0%; Pred. No. 2.3e+02;  
Matches 13; Conservative 2; Mismatches 8; Indels 3; Gaps 2;
```

```
QY   3 GTUWGSGMEAWLATVLKALPW-HPTY 27  
Db    | : | : | : |  
       ||| :||| :||| :||| :|||  
     19 GLLYGGGV--WAARVKNKHKEWLEPT 42
```

```
RESULT 15  
AAS75317 PRELIMINARY; PRT; 952 AA.
```

```
ID AAS75317  
AC AAS75317;  
DT 31-MAR-2004 (TEMBLrel. 27, Created)  
DI 31-MAR-2004 (TEMBLrel. 27, Last sequence update)  
DT 31-MAR-2004 (TEMBLrel. 27, Last annotation update)  
DE Calsynthenin.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OM Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Rattus.  
OX NCBI_TaxId=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Wistar;  
RA Ahmed F.; Torrado M., Zinovieva R.D., Senatorov V., Wistow G.,  
RA Tomarev S.I.;  
RT "Gene Expression Profile of the Rat Eye Irido-Corneal Angle. NEIbank  
RT Expressed Sequence Tag Analysis.";  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
RD ENBL; AY569014; AAS75317.1 - .  
SQ SEQUENCE 952 AA; 106260 MW; DBE2185AF20F9713 CRC64;  
  
Query Match      30.6%; Score 54.5; DB 2; Length 952;  
Best Local Similarity 50.0%; Pred. No. 2.3e+02;  
Matches 13; Conservative 2; Mismatches 8; Indels 3; Gaps 2;
```

```
QY   3 GTUWGSGMEAWLATVLKALPW-HPTY 27  
Db    | : | : | : |  
       ||| :||| :||| :||| :|||  
     19 GLLYGGGV--WAARVKNKHKEWLEPT 42
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 10.2491 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-60

Perfect score: 144

Sequence: 1 IAQATKATIDKWNCKIKIFYTSKKEAS 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA*

1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS-COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	50.7	60	4	US-09-513-999C-6532
2	71	49.3	107	4	US-09-513-999C-8172
3	48	33.3	43	4	US-09-270-767-37159
4	48	33.3	43	4	US-09-270-767-52376
5	47	32.6	732	4	US-09-270-767-38665
6	47	32.6	732	4	US-09-270-767-53882
7	46	31.9	278	4	US-09-270-767-39161
8	46	31.9	278	4	US-09-270-767-54378
9	45.5	31.6	134	4	US-09-107-532A-5878
10	45	31.2	198	4	US-09-134-000C-5033
11	45	31.2	250	4	US-09-270-767-52113
12	45	31.2	381	4	US-09-270-767-43720
13	45	31.2	650	4	US-09-270-767-42511
14	44	30.6	60	4	US-09-328-352-6326
15	44	30.6	476	4	US-09-248-796A-16327
16	43.5	30.2	190	1	US-08-106-981-2
17	43	29.9	119	4	US-09-107-532A-6608
18	42.5	29.5	159	4	US-09-270-767-34992
19	42.5	29.5	159	4	US-09-270-767-50209
20	42.5	29.5	177	4	US-09-270-767-37349
21	42.5	29.5	177	4	US-09-270-767-52566
22	42	29.2	192	4	US-09-319-588C-8
23	42	29.2	193	4	US-09-248-796A-15614
24	42	29.2	535	4	US-09-248-796A-18860
25	42	29.2	1024	4	US-09-562-737-41
26	42	29.2	2710	1	US-08-405-604A-6
27	42	29.2	2710	2	US-08-405-496A-6

28 42 29.2 2710 3 US-08-915-136-6 Sequence 6, Appli
29 42 29.2 2710 3 US-08-957-310-6 Sequence 6, Appli
30 42 29.2 2710 4 US-10-011-366-6 Sequence 6, Appli
31 42 29.2 2710 4 US-09-084-517-6 Sequence 6, Appli
32 41.5 28.8 286 4 US-09-134-000C-4791 Sequence 4791, Ap
33 41 28.5 37 4 US-09-270-767-61460 Sequence 61460, A
34 41 28.5 127 2 US-08-574-699A-2 Sequence 2, Appli
35 41 28.5 194 4 US-09-270-767-45922 Sequence 45922, A
36 41 28.5 198 4 US-09-270-767-32467 Sequence 32467, A
37 41 28.5 198 4 US-09-270-767-47684 Sequence 47684, A
38 41 28.5 265 4 US-09-248-796A-15796 Sequence 15796, A
39 41 28.5 268 4 US-09-371-338-21 Sequence 21, Appli
40 41 28.5 311 4 US-09-912-628-6 Sequence 6, Appli
41 41 28.5 494 4 US-09-252-991A-24183 Sequence 24183, A
42 41 28.5 690 4 US-09-371-338-19 Sequence 19, Appli
43 41 28.5 724 4 US-09-270-767-44066 Sequence 44066, A
44 41 28.5 1024 4 US-09-562-737-46 Sequence 46, Appli
45 41 28.5 1294 2 US-08-819-288-3 Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-513-999C-6532
; Sequence 6532, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 6532
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 49
; OTHER INFORMATION: Xaa-Ile or Lys
US-09-513-999C-6532

Query Match 50.7%; Score 73; DB 4; Length 60;
Best Local Similarity 62.5%; Pred.No. 0.00018;
Matches 15; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 AQATKATIDKWNCKIKIFYTSKK 25
Db 12 AMATKIDKWDLIKLSFCTAKE 35

RESULT 2
US-09-513-999C-8172
; Sequence 8172, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487

us-10-092-750-60.ra1

Fri Nov 12 14:55:47 2004

```
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 8172
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-8172

Query Match      49.3%; Score 71; DB 4; Length 107;
Best Local Similarity 62.5%; Pred. No. 0.00069;
Matches 15; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy  2 AQAATKATIDKWCNKIKLIFYTSKK 25
    |||||:|||||:|||||:|:|:|
Db  12 AWATKAKIDKWDLIKLSCTAKE 35

RESULT 3
US-09-270-767-37159
; Sequence 37159, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37159
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-37159

Query Match      33.3%; Score 48; DB 4; Length 43;
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy  8 TIDKWCNKIKLIFYTS 23
    ||:|:|:|:|:|:|:|
Db  1 TSDCYNVRLFIYSS 16

RESULT 4
US-09-270-767-52376
; Sequence 52376, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52376
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-52376

Query Match      33.3%; Score 48; DB 4; Length 43;
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy  8 TIDKWCNKIKLIFYTS 23
    ||:|:|:|:|:|:|:|
Db  1 TSDCYNVRLFIYSS 16

RESULT 5
US-09-270-767-52376
; Sequence 52376, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52376
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-52376

Query Match      33.3%; Score 48; DB 4; Length 43;
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy  8 TIDKWCNKIKLIFYTS 23
    ||:|:|:|:|:|:|:|
Db  1 TSDCYNVRLFIYSS 16

US-09-270-767-38665
; Sequence 38665, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38665
; LENGTH: 732
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-38665

Query Match      32.6%; Score 47; DB 4; Length 732;
Best Local Similarity 69.2%; Pred. No. 39;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy  13 NCIKLKIFYTSKK 25
    |||||:|:|:|
Db  588 NCIKLLLYRSIK 600

RESULT 6
US-09-270-767-53882
; Sequence 53882, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53882
; LENGTH: 732
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-53882

Query Match      32.6%; Score 47; DB 4; Length 732;
Best Local Similarity 69.2%; Pred. No. 39;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy  13 NCIKLKIFYTSKK 25
    |||||:|:|:|
Db  588 NCIKLLLYRSIK 600

RESULT 7
US-09-270-767-39161
; Sequence 39161, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39161
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-39161
```

```
;
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-39161

Query Match      31.9%; Score 46; DB 4; Length 278;
Best Local Similarity 40.9%; Pred. No. 20;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 3 QATKATIDKWNCKIKLIFYTSK 24
   ||| ||| :|: |||
Db 188 QATPLTICSASCLDFRLVITSK 209

RESULT 8
US-09-270-767-54378
; Sequence 54378, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54378
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-54378

Query Match      31.9%; Score 46; DB 4; Length 278;
Best Local Similarity 40.9%; Pred. No. 20;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 3 QATKATIDKWNCKIKLIFYTSK 24
   ||| ||| :|: |||
Db 188 QATPLTICSASCLDFRLVITSK 209

RESULT 9
US-09-107-532A-5878
; Sequence 5878, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Denéke
```

```
;
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5878:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 134 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (B) LOCATION 1...134
; SEQUENCE DESCRIPTION: SEQ ID NO: 5878:
US-09-107-532A-5878

Query Match      31.6%; Score 45.5; DB 4; Length 134;
Best Local Similarity 47.6%; Pred. No. 10;
Matches 10; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

Qy 9 IDKNCIKLKI-FYTSKEAS 28
   ||| ||| ||| |||
Db 2 IDRWATSKIKIRFYSEDSGS 22

RESULT 10
US-09-134-000C-5033
; Sequence 5033, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5033
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5033

Query Match      31.2%; Score 45; DB 4; Length 198;
Best Local Similarity 61.5%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IAQATKATIDKWN 13
   :|: ||| |||
Db 82 LADTLKATIDTWN 94

RESULT 11
US-09-270-767-59113
; Sequence 59113, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59113
; LENGTH: 250
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 32.3849 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-60

Perfect score: 144

Sequence: 1 IAQATKATIDKWNCKIKIFYTSKKEAS 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	144	100.0	28	14	US-10-092-750-60
2	79	54.9	317	13	US-10-083-853-2
3	79	54.9	464	15	US-10-211-462-221
4	79	54.9	641	16	US-10-408-765A-181
5	79	54.9	843	15	US-10-001-885-115
6	79	54.9	917	13	US-10-001-835-189
7	79	54.9	940	14	US-10-082-830-235
8	79	54.9	1010	14	US-10-050-882-80
9	79	54.9	1031	14	US-10-082-828A-257
10	79	54.9	1081	15	US-10-085-198-150
11	79	54.9	1272	15	US-10-114-270-30
12	79	54.9	1274	16	US-10-415-615-3
13	78	54.2	219	14	US-10-094-749-1795

```

14 53.5 101 9 US-09-867-550-1160 Sequence 1160, Ap
15 52.8 647 14 US-10-021-660-94 Sequence 94, Appl
16 52.1 990 15 US-10-085-198-154 Sequence 154, App
17 50.7 1018 15 US-10-085-198-156 Sequence 156, App
18 50.7 1275 14 US-10-025-201-3 Sequence 3, Appli
19 50.0 126 15 US-10-108-260A-4667 Sequence 4667, Ap
20 50.0 956 14 US-10-000-256A-153 Sequence 153, App
21 49.3 123 14 US-10-104-047-3008 Sequence 2008, Ap
22 49.3 206 14 US-10-104-047-3744 Sequence 3744, Ap
23 48.6 146 15 US-10-108-260A-2729 Sequence 2729, Ap
24 46.5 1017 15 US-10-085-198-152 Sequence 152, App
25 37.5 48 9 US-09-925-300-1097 Sequence 1097, Ap
26 37.5 81 15 US-10-276-774-1716 Sequence 1716, Ap
27 35.4 103 15 US-10-276-774-2492 Sequence 2492, Ap
28 34.7 120 15 US-10-424-599-178583 Sequence 178583,
29 34.0 1013 15 US-10-092-900A-230 Sequence 230, App
30 34.0 1013 15 US-10-336-472-128 Sequence 128, App
31 34.0 1013 15 US-10-236-472-126 Sequence 126, App
32 34.0 1018 15 US-10-336-472-126 Sequence 126, App
33 34.0 1018 15 US-10-236-472-126 Sequence 126, App
34 33.3 49 15 US-10-424-599-188449 Sequence 188449,
35 46.5 111 16 US-10-767-701-44455 Sequence 44455, A
36 46.5 186 17 US-10-425-115-280905 Sequence 280905,
37 46 87 17 US-10-425-115-233963 Sequence 233963,
38 46 373 16 US-10-479-284-30 Sequence 30, Appl
39 46 1842 14 US-10-243-552-847 Sequence 847, App
40 46 1842 14 US-10-243-552-887 Sequence 887, App
41 45.5 94 15 US-10-335-977-8018 Sequence 8018, Ap
42 45.5 127 15 US-10-335-977-8019 Sequence 8019, Ap
43 45.5 497 9 US-09-981-353-105 Sequence 105, App
44 45.5 508 14 US-10-170-385-347 Sequence 347, App
45 45 110 17 US-10-425-115-246081 Sequence 246081,

```

ALIGNMENTS

```

RESULT 1
US-10-092-750-60
; Sequence 60, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-60
Query Match 100.0%; Score 144; DB 14; Length 28;
Best Local Similarity 100.0%; Pred No. 1.8e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IAQATKATIDKWNCKIKIFYTSKKEAS 28
Db 1 IAQATKATIDKWNCKIKIFYTSKKEAS 28

```

```

RESULT 2
US-10-083-853-2
; Sequence 2, Application US/10083853
; Publication No. US20020164709A1
; GENERAL INFORMATION:

```

; APPLICANT: Affymetrix, Inc
 ; APPLICANT: Shigeta, Ron T
 ; APPLICANT: Siani-Rose, Michael A
 ; TITLE OF INVENTION: Nucleic Acid Encoding Growth Factor Protein
 ; FILE REFERENCE: 3385.1
 ; CURRENT APPLICATION NUMBER: US/10/083,853
 ; PRIOR FILING DATE: 2002-02-26
 ; PRIOR APPLICATION NUMBER: USSN 60/272,663
 ; PRIOR FILING DATE: 2001-03-01
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 317
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-10-083-853-2

Query Match 54.9%; Score 79; DB 13; Length 317;
 Best Local Similarity 66.7%; Pred. No. 0.0014;
 Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 A0ATKATIDKWCNCKIKIFYTSKK 25
 Db 12 AMATKAKIDKWDLIKLSFCTAKE 35

RESULT 3

US-10-211-462-221
 ; Sequence 221, Application US/10211462
 ; Publication No. US2004003495A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Murray, Richard
 ; APPLICANT: Glynn, Richard
 ; APPLICANT: Watson, Susan R.
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
 ; TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
 ; FILE REFERENCE: 018501-006200US
 ; CURRENT APPLICATION NUMBER: US/10/211,462
 ; CURRENT FILING DATE: 2003-02-13
 ; PRIOR APPLICATION NUMBER: US 09/784,356
 ; PRIOR FILING DATE: 2001-02-14
 ; PRIOR APPLICATION NUMBER: US 09/791,390
 ; PRIOR FILING DATE: 2001-02-22
 ; PRIOR APPLICATION NUMBER: US 60/310,025
 ; PRIOR FILING DATE: 2001-08-03
 ; PRIOR APPLICATION NUMBER: US 60/334,244
 ; PRIOR FILING DATE: 2001-11-29
 ; NUMBER OF SEQ ID NOS: 230
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 221
 ; LENGTH: 464
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-211-462-221

Query Match 54.9%; Score 79; DB 15; Length 464;
 Best Local Similarity 66.7%; Pred. No. 0.0021;
 Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 A0ATKATIDKWCNCKIKIFYTSKK 25
 Db 12 AMATKAKIDKWDLIKLSFCTAKE 35

RESULT 4

US-10-408-765A-181
 ; Sequence 181, Application US/10408765A
 ; Publication No. US2004010187A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ghosh, Soumitra S.
 ; APPLICANT: Fahy, Eoin D.

; APPLICANT: Zhang, Bing
 ; APPLICANT: Gibson, Bradford W.
 ; APPLICANT: Taylor, Steven W.
 ; APPLICANT: Glenn, Gary M.
 ; APPLICANT: Warnock, Dale E.
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 ; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
 ; FILE REFERENCE: 650088.465
 ; CURRENT APPLICATION NUMBER: US/10/408,765A
 ; CURRENT FILING DATE: 2003-04-04
 ; NUMBER OF SEQ ID NOS: 3077
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 181
 ; LENGTH: 641
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-408-765A-181

Query Match 54.9%; Score 79; DB 16; Length 641;
 Best Local Similarity 66.7%; Pred. No. 0.0023;
 Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 A0ATKATIDKWCNCKIKIFYTSKK 25
 Db 367 AMATKAKIDKWDLIKLSFCTAKE 390

RESULT 5

US-10-001-885-115
 ; Sequence 115, Application US/10001885
 ; Publication No. US20040058319A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Salceda, Susana
 ; APPLICANT: Macina, Roberto
 ; APPLICANT: Recipon, Herve
 ; APPLICANT: Cafferkey, Robert
 ; APPLICANT: Sun, Yongming
 ; APPLICANT: Liu, Chenghua
 ; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pr
 ; FILE REFERENCE: DEX-0279
 ; CURRENT APPLICATION NUMBER: US/10/001,885
 ; CURRENT FILING DATE: 2001-11-20
 ; PRIOR APPLICATION NUMBER: 60/252,061
 ; PRIOR FILING DATE: 2000-11-20
 ; PRIOR APPLICATION NUMBER: 60/253,257
 ; PRIOR FILING DATE: 2000-11-27
 ; NUMBER OF SEQ ID NOS: 167
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 115
 ; LENGTH: 843
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 US-10-001-885-115

Query Match 54.9%; Score 79; DB 15; Length 843;
 Best Local Similarity 66.7%; Pred. No. 0.0039;
 Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 A0ATKATIDKWCNCKIKIFYTSKK 25
 Db 715 AMATKAKIDKWDLIKLSFCTAKE 738

RESULT 6

US-10-001-835-189
 ; Sequence 189, Application US/10001835
 ; Publication No. US20020160387A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Salceda, Susana
 ; APPLICANT: Macina, Roberto
 ; APPLICANT: Recipon, Herve
 ; APPLICANT: Cafferkey, Robert
 ; APPLICANT: Sun, Yongming


```
; APPLICANT: Liu, Chenghua
; FILE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro
; FILE REFERENCE: DEX-0277
; CURRENT APPLICATION NUMBER: US/10/001,835
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/249,997
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 189
; TYPE: PRT
; LENGTH: 917
; ORGANISM: Homo sapien
US-10-001-835-189

Query Match      54.9%; Score 79; DB 13; Length 917;
Best Local Similarity 56.7%; Pred. No. 0.0043;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      2 AOAATKATIDKWNCKIKLKFVTSKK 25
DB      790 AATKAKIDKNDLKLKSFCTAKE 813

RESULT 7
US-10-082-830-235
; Sequence 235, Application US/10082830
; Publication No. US2003007604A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Salceda, Susana
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
; TITLE OF INVENTION: Genes and Proteins
; FILE REFERENCE: DEX-0249
; CURRENT APPLICATION NUMBER: US/10/082,830
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,802
; PRIOR FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 282
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 235
; TYPE: PRT
; LENGTH: 940
; ORGANISM: Homo sapiens
US-10-082-830-235

Query Match      54.9%; Score 79; DB 14; Length 940;
Best Local Similarity 66.7%; Pred. No. 0.0044;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      2 AOAATKATIDKWNCKIKLKFVTSKK 25
DB      812 AATKAKIDKNDLKLKSFCTAKE 835

RESULT 8
US-10-050-882-80
; Sequence 80, Application US/10050882
; Publication No. US20030104400A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 27 Human secreted proteins
; FILE REFERENCE: P2038P1
; CURRENT APPLICATION NUMBER: US/10/050,882
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/661,453
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: PCT/US00/06783
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/125,055

; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 80
; LENGTH: 1010
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (25)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (104)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (194)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (362)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (525)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (643)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (649)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (656)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (660)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (731)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (770)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (777)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (790)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (800)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (825)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (987)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (996)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (1003)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-050-882-80

Query Match      54.9%; Score 79; DB 14; Length 1010;
Best Local Similarity 66.7%; Pred. No. 0.0047;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      2 AOAATKATIDKWNCKIKLKFVTSKK 25
DB      736 AATKAKIDKNDLKLKSFCTAKE 759
```

```

RESULT 9
US-10-082-828A-257
; Sequence 257, Application US/10082828A
; Publication No. US2003017515A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Salceda, Susana
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and P
; FILE REFERENCE: DEX-0247
; CURRENT APPLICATION NUMBER: US/10/082,828A
; CURRENT FILING DATE: 2002-07-09
; PRIOR APPLICATION NUMBER: US 60/243,805
; PRIOR FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 257
; LENGTH: 1031
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-082-828A-257

Query Match          54.9%; Score 79; DB 14; Length 1031;
Best Local Similarity 66.7%; Pred. No. 0.0048;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY  2 AOAATKATIDKWCICIKLIFYTSKK 25
    |||||:||||:||||:||||:
DB  903 AMATKAKIDKWDLIKLSFCTAKE 926

RESULT 10
US-10-085-198-150
; Sequence 150, Application US/10085198
; Publication No. US2004009907A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-279
; CURRENT APPLICATION NUMBER: US/10/085,198
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/271,646
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/276,401
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,981
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/312,858
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/271,840
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/277,324
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/286,096
; PRIOR FILING DATE: 2001-04-21
; PRIOR APPLICATION NUMBER: 60/299,695
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/315,614
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/272,405
; PRIOR FILING DATE: 2001-02-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 653
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 150
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-085-198-150

Query Match          54.9%; Score 79; DB 15; Length 1081;
Best Local Similarity 66.7%; Pred. No. 0.005;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY  2 AOAATKATIDKWCICIKLIFYTSKK 25
    |||||:||||:||||:||||:
DB  782 AMATKAKIDKWDLIKLSFCTAKE 805

RESULT 11
US-10-114-270-30
; Sequence 30, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Ziaohong
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glenna
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Gangoli, Bsha A.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Ji, Weizhen
; APPLICANT: Anderson, David W.
; APPLICANT: Liette, Mario W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: MacDougall, John R.
; APPLICANT: Rothenberg, Mark E.
; TITLE OF INVENTION: NO. US20040030110A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-322C
; CURRENT APPLICATION NUMBER: US/10/114,270
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 30
; LENGTH: 1272
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-085-198-150
```

US-10-114-270-30

Query Match 54.9%; Score 79; DB 15; Length 1272;
Best Local Similarity 66.7%; Pred. No. 0.0059;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 AQAATKATIDKWCNKIKLIFYTSKK 25
DB 992 AMATKAKIDKWDLIKLSFKTAK 1015

RESULT 12

US-10-415-615-3

; Sequence 3, Application US/10415615

; Publication No. US20040101943A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE CORPORATION

; APPLICANT: BAUGHN, Mariah R.

; APPLICANT: YUE, Henry

; APPLICANT: LU, Yan

; APPLICANT: DING, Li

; APPLICANT: TANG, Y. Tom

; APPLICANT: GANDHI, Ameena R.

; APPLICANT: RAFALIA, April J.A.

; APPLICANT: LAL, Preeti G.

; TITLE OF INVENTION: NUCLEIC ACID MODIFICATION ENZYMES

; FILE REFERENCE: PI-0280 USN

; CURRENT APPLICATION NUMBER: US/10/415,615

; CURRENT FILING DATE: 2003-04-29

; PRIOR APPLICATION NUMBER: PCT/US01/46301

; PRIOR FILING DATE: 2001-11-01

; PRIOR APPLICATION NUMBER: US 60/245,458

; PRIOR FILING DATE: 2000-11-02

; PRIOR APPLICATION NUMBER: US 60/255,107

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PERL Program

; SEQ ID NO 3

; LENGTH: 1274

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No: 1258887CD1

US-10-415-615-3

Query Match 54.9%; Score 79; DB 16; Length 1274;
Best Local Similarity 66.7%; Pred. No. 0.0059;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 AQAATKATIDKWCNKIKLIFYTSKK 25
DB 1000 AMATKAKIDKWDLIKLSFKTAK 1023

RESULT 13

US-10-094-749-1795

; Sequence 1795, Application US/10094749

; Publication No. US20030219741A1

; GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: OTSUKI, TETSUJI

; APPLICANT: WAKAMATSU, AI

; APPLICANT: SATO, HIROYUKI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: YAMAMOTO, JUN-ICHI

; APPLICANT: ISONO, YUUKO

; APPLICANT: HIO, YURI

; APPLICANT: OTSUKA, KAORU

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: IRIE, EVOTARO

; APPLICANT: TAMECHIKA, ICHIRO

; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1795
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-1795

Query Match 54.2%; Score 78; DB 14; Length 219;
Best Local Similarity 63.0%; Pred. No. 0.0014;
Matches 17; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 AQAATKATIDKWCNKIKLIFYTSKKEAS 28
DB 67 AMATKAKIDKWDLIKLSFGTAKDTIS 93

RESULT 14

US-09-867-550-1160

; Sequence 1160, Application US/09867550

; Patent No. US20020082206A1

; GENERAL INFORMATION:

; APPLICANT: Leach, Martin D.

; APPLICANT: Mehraban, Fuad,

; APPLICANT: Conley, Pamela

; APPLICANT: Law, Debbie

; APPLICANT: Topper, James

; TITLE OF INVENTION: NO. US20020082206A1 Polynucleotides from Atherogenic Cells and

; TITLE OF INVENTION: Thereby

; FILE REFERENCE: 21402-013 (Cura-313)

; CURRENT APPLICATION NUMBER: US/09/867,550

; CURRENT FILING DATE: 2001-09-20

; PRIOR APPLICATION NUMBER: USSN 60/208,427

; PRIOR FILING DATE: 2000-05-30

; NUMBER OF SEQ ID NOS: 2125

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1160

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (3)

; OTHER INFORMATION: Wherein Xaa may be any one of Ala or Asp or Gly or Val

US-09-867-550-1160

Query Match 53.5%; Score 77; DB 9; Length 101;
Best Local Similarity 58.3%; Pred. No. 0.00089;
Matches 14; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 AQAATKATIDKWCNKIKLIFYTSKK 25
DB 15 AEATKAKLDNWDICIKLNFCMAKE 38

RESULT 15

US-10-021-660-94

; Sequence 94, Application US/10021660

; Publication No. US20030152926A1

; GENERAL INFORMATION:

```

; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152926A1el Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 94
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-021-660-94

```

```

Query Match      52.8%; Score 76; DB 14; Length 647;
Best Local Similarity 66.7%; Pred. No. 0.0084;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      2 AQAATKATIDKWCICIKLIFYTSKK 25
      ||||| ||||| ||||| |||||
Db      12 AFATKAKIDKWDLIKLSFCTAKE 35

```

Search completed: November 11, 2004, 02:43:16
Job time : 33.4349 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 6.86792 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-60
Perfect score: 144
Sequence: 1 IAQATKATIDKWNCKIKLIFYTSKKEAS 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416.

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	54.9	562	2 JU0033	hypothetical L1 pr
2	79	54.9	712	2 S23650	retrovirus-related
3	79	54.9	1275	2 B28096	line-1 protein ORF
4	79	54.9	1280	2 B34087	hypothetical prote
5	73	50.7	1275	2 I38588	reverse transcript
6	73	50.7	1275	2 S85824	reverse transcript
7	72	50.0	1259	4 GNRUL1	retrovirus-related
8	69	47.9	92	2 F41925	hypothetical prote
9	56	38.9	275	2 S21348	probable pol polyp
10	56	38.9	513	2 S21976	probable RNA-direc
11	52	36.1	1281	1 GMSLL	retrovirus-related
12	50	34.7	184	2 T38315	very hypothetical
13	50	34.7	806	2 F69899	phage-related pre-
14	49	34.0	488	1 OXASBI	mRNA maturase bil
15	48	33.3	354	2 A52138	hypothetical prote
16	48	33.3	392	2 S57662	hypothetical prote
17	47.5	33.0	545	2 D90159	hypothetical prote
18	47.5	33.0	2712	2 T30949	hypothetical prote
19	47.5	33.0	3864	2 D37757	protein C44E4.1a [
20	47	32.6	206	2 JC2574	hypothetical 22k p
21	46	31.9	1260	4 GNLRL1	retrovirus-related
22	45.5	31.6	94	2 B64550	hypothetical prote
23	45.5	31.6	94	2 F71959	hypothetical prote
24	45.5	31.6	508	1 S62599	UTP-glucose-1-phos
25	45.5	31.6	508	1 JX0277	UTP-glucose-1-phos
26	45.5	31.6	508	2 S35692	UTP-glucose-1-phos
27	45	31.2	129	2 D64444	hypothetical prote
28	45	31.2	485	2 G97324	probable membrane
29	45	31.2	494	2 T05302	vacuolar processin

30	45	31.2	550	2 S64725	probable lipoprote
31	45	31.2	892	2 H71426	hypothetical prote
32	45	31.2	1309	2 F82207	ATP-dependent heli
33	44.5	30.9	1437	2 S07430	M polypeptide prec
34	44	30.6	128	2 E96790	hypothetical prote
35	44	30.6	334	2 D86455	T9L6.3 protein - A
36	44	30.6	356	2 E84495	En/Spm-like transp
37	44	30.6	365	2 B84515	hypothetical prote
38	44	30.6	423	2 T03119	hypothetical prote
39	44	30.6	436	2 B85469	protein Flxk21.16
40	44	30.6	437	2 G84505	probable PttA2-lik
41	44	30.6	446	2 F84506	En/Spm-like transp
42	44	30.6	459	2 G84752	En/Spm-like transp
43	44	30.6	550	2 B82912	arginyl tRNA synth
44	44	30.6	756	2 S60966	probable protein k
45	44	30.6	1893	2 A56158	eye development pr

ALIGNMENTS

RESULT 1

Hypothetical L1 protein (third intron of gene TS) - human
C:Species: Homo sapiens (man)
C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C:Accession: JU0033
R:Horie, N.; Nalbantoglu, J.; Kaneda, S.; Ayusawa, D.; Seno, T.; Takeishi, K.
J. Biochem. 106: 1-4, 1989
A>Title: Identification and characterization of an L1 family sequence with a very long
A:Reference number: JU0033; MUID:89380111; PMID:2476429
A:Status: nucleic acid sequence not shown
A:Accession: JU0033
A:Molecule type: DNA
A:Residues: 1-562 <HOR>
A:Cross-references: UNIPROT:Q00378
A>Note: this sequence is similar to human teratocarcinoma L1 RNA species and RNA depend
C:Superfamily: pol polyprotein

Query Match 54.9%; Score 79; DB 2; Length 562;
Best Local Similarity 66.7%; Pred. No. 0.00048;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy	2	AQATKATIDKWNCKIKLIFYTSKK 25
Db	489	AMATKAKIDKMDLKLKSFCTAKE 512

RESULT 2

S23650
retrovirus-related hypothetical protein II - human retrotransposon LINE-1
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004
C:Accession: S23650
R:Hohjoh, H.; Minakami, R.; Sakaki, Y.
Nucleic Acids Res. 18, 4099-4104, 1990
A>Title: Selective cloning and sequence analysis of the human L1 (LINE-1) sequences whi
A:Reference number: S23649; MUID:90332398; PMID:2165587
A:Accession: S23650
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-712 <HOH>
A:Cross-references: UNIPROT:Q15604; UNIPROT:Q12881; UNIPROT:Q00363; UNI
PROT:Q14754; UNIPROT:Q915K0; UNIPROT:Q8E30; UNIPROT:Q00375; EMBL:X522;
C:Genetics:
A>Note: the nucleotide sequence was submitted to the ENBL Data Library, March 1990
A:Mobile element: LINE-1
A:Start codon: GTG
C:Superfamily: pol polyprotein

Query Match 54.9%; Score 79; DB 2; Length 712;
Best Local Similarity 66.7%; Pred. No. 0.0006;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

A:Accession: J98315
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-184 <BRO>
A:Cross-references: UNIPROT:O13954; EMBL:Z98977; PIDN:CAB11667.1; GSPDB:GN000666
A:Experimental source: strain 972h; cosmid c23H4
C:Genetics:
A:Gene: SPDB:SPAC23H4.13C
A:Map position: 1
A:Introns: 57/3
C:Superfamily: Schizosaccharomyces pombe very hypothetical protein SPAC23H4.13C

Query Match 34.7%; Score 50; DB 2; Length 184;
 Best Local Similarity 66.7%; Pred. No. 4.2; Mismatches 3; Indels 1; Gaps 0;
 Matches 8; Conservative 3; Mismatches 3; Indels 1; Gaps 0;

QY 8 TIDKWNCKIKKI 19
 |||:|||||:
 Db 104 TIEKWSCKIEKL 115

RESULT 13
 F69899
 phage-related pre-neck appendage protein homolog yobo - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: F69899
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallex
 ieck, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A:Authors: Tauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
 Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
 akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terptrata, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumstein, B.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:198044033; PMID:9384377
 A:Accession: F69899
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-806 <KUN>
 A:Cross-references: UNIPROT:Q34433; GB:Z99114; GB:AL009126; NID:92634230; PIDN:CAB13795.
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yobo
 C:Superfamily: phage PZA gene 12 protein

Query Match 34.7%; Score 50; DB 2; Length 806;
 Best Local Similarity 47.6%; Pred. No. 17;
 Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 5 TKATIDKWNCKIKKIFYSKK 25
 |||:|||||:
 Db 776 TKATEGSKWKMKITSAYTSEK 796

RESULT 14
 QXASBI
 mRNA maturase bil - Emericella nidulans mitochondrion
 N:Alternate names: coba intron protein
 C:Species: mitochondrion Emericella nidulans, Aspergillus nidulans
 C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 09-Jul-2004
 C:Accession: A04513; A37516
 R:Waring, R.B.; Davies, R.W.; Lee, S.; Grisi, E.; Berks, M.M.; Scazzocchio, C.
 Cell 27, 4-11, 1981
 A:Title: The mosaic organization of the apocytochrome b gene of Aspergillus nidulans rev
 A:Reference number: A00157; MUID:82115341; PMID:7034966
 A:Accession: A04513
 A:Molecule type: DNA
 A:Residues: 1-237 <WAL>
 A:Cross-references: UNIPROT:P03880; GB:J01388; GB:V00651; GB:V00652; NID:g336
 A:Experimental source: imperfect stage
 A:Note: This ORF is not annotated in GenBank entry EMENTCYB2, release 106.0
 R:Waring, R.B.; Davies, R.W.; Scazzocchio, C.; Brown, T.A.
 Proc. Natl. Acad. Sci. U.S.A. 79, 6332-6336, 1982
 A:Title: Internal structure of a mitochondrial intron of Aspergillus nidulans.
 A:Reference number: A37516; MUID:83065170; PMID:6755468
 A:Accession: A37516
 A:Molecule type: DNA

A:Residues: 161-488 <WA2>
 A:Cross-references: GB:J01387; NID:g336899
 A:Experimental source: imperfect stage
 A:Note: this ORF is not annotated in GenBank entry EMENTCOBA, release 106.0
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC3
 C:Superfamily: Emericella mRNA maturase bil; cytochrome b homology; cytochrome b6 homol
 C:Keywords: mitochondrion
 F:1-168/Region: cytochrome b exon encoded
 F:10-168/Domain: cytochrome b homology (fragment) <CBH>
 F:10-168/Domain: cytochrome b6 homology (fragment) <CB6>
 F:169-488/Region: cytochrome b intron encoded

Query Match 34.0%; Score 49; DB 1; Length 488;
 Best Local Similarity 60.0%; Pred. No. 15;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 TKATIDKWNCKIKKI 19
 |||:|||||:
 Db 423 TKVYLDKTNCKSLKV 437

RESULT 15
 AE2138
 hypothetical protein alr2660 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C:Accession: AE2138
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AE2138
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-354 <KUR>
 A:Cross-references: UNIPROT:Q9YQT3; GB:BA000019; PIDN:BA074359.1; PID:gl7131753; GSPDB:
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr2660
 C:Superfamily: uncharacterized protein with thioresdoxin-like domain

Query Match 33.3%; Score 48; DB 2; Length 354;
 Best Local Similarity 40.0%; Pred. No. 16;
 Matches 10; Conservative 5; Mismatches 6; Indels 4; Gaps 1;

QY 6 KATIDKWN---CIKLKIFYSKKE 26
 |||:|||||:
 Db 289 KLLVERNEQATWVKLAISYTSKDD 313

Search completed: November 10, 2004, 14:52:27
 Job time : 7.86792 secs

RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
 RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
 RA Isogai T.,
 RA Submitted (JUL-2003) to the EMBL/GenBank/DBSJ databases.
 DR EMBL; AK125948; BAC6356.1; --
 SQ SEQUENCE 254 AA; 30123 MW; AFE99F7AB72FA49D CRC64;

Query Match 54.9%; Score 79; DB 2; Length 254;
 Best Local Similarity 66.7%; Pred. No. 0.00075;
 Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 AOATKATIDKWNCKIKLIFYTSKK 25
 DB 12 AMATKAKIDKWLKIKSFTAKE 35

RESULT 3
 ID BAC6356 PRELIMINARY; PRT; 254 AA.
 AC BAC6356;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE CDNA FLJ43960 fis, clone T5ST14016551.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primata; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Testis;
 RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
 RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
 RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
 RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
 RA Isogai T.;
 RA "NEDO human cDNA sequencing project.";
 RT Submitted (JUL-2003) to the EMBL/GenBank/DBSJ databases.
 DR EMBL; AK125948; BAC6356.1; --
 SQ SEQUENCE 254 AA; 30123 MW; AFE99F7AB72FA49D CRC64;

Query Match 54.9%; Score 79; DB 2; Length 254;
 Best Local Similarity 66.7%; Pred. No. 0.00075;
 Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 AOATKATIDKWNCKIKLIFYTSKK 25
 DB 12 AMATKAKIDKWLKIKSFTAKE 35

RESULT 4
 ID Q8NB08 PRELIMINARY; PRT; 473 AA.
 AC Q8NB08;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein FLJ34421.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC PubMed=14702039;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
 RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,

RA Sudo H., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa B., Omura Y.,
 RA Abe K., Kanihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
 RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa M., Fujimori K.,
 RA Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y., Ishida S.,
 RA Ono Y., Takiguchi S., Watanabe S., Yoisida M., Hoshino T., Kusano J.,
 RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
 RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
 RA Yoshino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshihawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Moniyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigetani K., Senba T.,
 RA Katsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu I.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 DR EMBL; AK091740; BAC03736.1; --
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR InterPro: IPR000477; RVTse.
 DR Pfam: PF00078; RVT; 1.
 DR RNA-directed DNA polymerase; Transferase.
 KW RNA-directed DNA polymerase; 55258 MW; 5DD23CF8849FABDA CRC64;
 SQ SEQUENCE 473 AA; 55258 MW; 5DD23CF8849FABDA CRC64;

Query Match 54.9%; Score 79; DB 2; Length 473;
 Best Local Similarity 66.7%; Pred. No. 0.0014;
 Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 AOATKATIDKWNCKIKLIFYTSKK 25
 DB 400 AMATKAKIDKWLKIKSFTAKE 423

RESULT 5
 ID Q14754 PRELIMINARY; PRT; 712 AA.
 AC Q14754;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE ORF1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Japanese; TISSUE=Placenta;
 RX MEDLINE=90332398; PubMed=21655587;
 RA Honjoh H., Minakami R., Sakaki Y.;
 RT "Selective cloning of the Human LI (line-1) sequence which transposed
 RT in a relatively recent past.";
 RL Nucleic Acids Res. 18:4099-4104(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Japanese; TISSUE=Placenta;
 RX MEDLINE=92319645; PubMed=1320255;
 RA Minakami R., Kurase K., Etoh K., Furuhata Y., Hattori M., Sakaki Y.;
 RT "Identification of an internal cis-element essential for the human LI
 RT transcription and a nuclear factor(s) binding to the element.";
 RL Nucleic Acids Res. 20:3139-3145(1992).

```

RN RP SEQUENCE FROM N.A.
RN MEDLINE=97285120; PubMed=9140393;
RA Sassaman D.M., Dombroski B.A., Moran J.V., Kimberland M.L., Naas T.P.,
RA DeBerardinis R.J., Gabriel A., Swergold G.D., Kazazian H.H. Jr.;
RT "Many human L1 elements are capable of retrotransposition.";
RL Nat. Genet. 16:37-43(1997).
DR EMBL: U93564; AAC51263.1; -.
DR PIR: B28096; B28096.
DR PIR: S23650; S23650.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro: IPR005135; Exo_endo_phos.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF03372; Exo_endo_phos; 1.
DR Pfam: PF00078; RVT; 1.
DR RNA-directed DNA polymerase; Transferrase.
SQ SEQUENCE 1275 AA; 149201 MW; 23D516D624358F28 CRC64;

Query Match 54.9%; Score 79; DB 2; Length 1275;
Best Local Similarity 66.7%; Pred.No. 0.0038;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 AQATKATIDKWCNCKLKIFVTSKK 25
DB 1001 AMATKAKIDKWDLIKLSFCTAKE 1024
||||| ||||| :||| :|:

RESULT 8
OO0372 PRELIMINARY; PRT; 1275 AA.
AC OO0372;
DT 01-JUL-1997 (T-EMBLrel. 04, Created)
DT DT 01-JUL-1997 (T-EMBLrel. 04, Last sequence update)
DT DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Putative p150.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
XX [1];
RN RP SEQUENCE FROM N.A.
RN MEDLINE=97285120; PubMed=9140393;
RA Sassaman D.M., Dombroski B.A., Moran J.V., Kimberland M.L., Naas T.P.,
RA DeBerardinis R.J., Gabriel A., Swergold G.D., Kazazian H.H. Jr.;
RT "Many human L1 elements are capable of retrotransposition.";
RL Nat. Genet. 16:37-43(1997).
DR EMBL: U93570; AAC51273.1; -.
DR PIR: B28096; B28096.
DR PIR: S23650; S23650.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro: IPR005135; Exo_endo_phos.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF03372; Exo_endo_phos; 1.
DR Pfam: PF00078; RVT; 1.
DR RNA-directed DNA polymerase; Transferrase.
SQ SEQUENCE 1275 AA; 149110 MW; 4711B3BC22F7674E CRC64;

Query Match 54.9%; Score 79; DB 2; Length 1275;
Best Local Similarity 66.7%; Pred.No. 0.0038;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 AQATKATIDKWCNCKLKIFVTSKK 25
DB 1001 AMATKAKIDKWDLIKLSFCTAKE 1024
||||| ||||| :||| :|:

RESULT 9
OO0375 PRELIMINARY; PRT; 1275 AA.
ID OO0375

```

Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0

QY 2 A0ATKATIDKWNCKIKLKFVTSKK 25
 DB 1001 AMATKAKIDKWLKLSFCTAKE 1024

RESULT 11
 Q6ZVR1
 ID Q6ZVR1 PRELIMINARY; PRT; 164 AA.
 AC Q6ZVR1
 DT 05-JUL-2004 (TReMBLrel. 27, Created)
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
 DE Hypothetical protein FLJ42200.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thymus;
 RA Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Sugiyama T.,
 RA Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J.,
 RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
 RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
 RA Masuho Y., Nagai K., Isogai T.,
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK124194; BAC85800.1; -
 SQ SEQUENCE 164 AA; 19705 MW; 049A53C2DCCF46A6 CRC64;

Query Match 52.1%; Score 75; DB 2; Length 164;
 Best Local Similarity 68.2%; Pred. No. 0.002;
 Matches 15; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATKATIDKWNCKIKLKFVTSKK 25
 DB 9 ATKAKIDKWLKLSFCTAKE 30

RESULT 12
 BAC85800
 ID BAC85800 PRELIMINARY; PRT; 164 AA.
 AC BAC85800;
 DT 02-MAR-2004 (TReMBLrel. 27, Created)
 DT 02-MAR-2004 (TReMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TReMBLrel. 27, Last annotation update)
 DE CDNA FLJ42200 fis, clone THYMU2034647.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thymus;
 RA Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Sugiyama T.,
 RA Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J.,
 RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
 RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
 RA Masuho Y., Nagai K., Isogai T.,
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK124194; BAC85800.1; -
 SQ SEQUENCE 164 AA; 19705 MW; 049A53C2DCCF46A6 CRC64;

Query Match 52.1%; Score 75; DB 2; Length 164;
 Best Local Similarity 68.2%; Pred. No. 0.002;
 Matches 15; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATKATIDKWNCKIKLIFYTSKK 25
 |||||:|||||:|:|:
 Db 9 ATKAKIDKWDLIKLSFCTAKE 30

RESULT 13

QSTES0
 ID Q8TE30 PRELIMINARY; PRT; 1275 AA.
 AC Q8TE30;
 DT 01-JUN-2002 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21668188; PubMed=11810275;
 RA Benjes S.M., Morris C.M.;
 RT "A full-length and potentially active LINE element is integrated
 RT polymorphically within the IGL locus in a genomically unstable region
 RT of chromosome 22.";
 RL Hum. Genet. 109:628-637(2001).
 DR EMBL; AF421375; AAL50637.1; -;
 DR PIR; B28096; B28096.
 DR PIR; S23650; S23650.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR InterPro; IPR005135; Exo_endo_phos.
 DR InterPro; IPR000477; RVTse.
 DR Pfam; PF03372; Exo_endo_phos; 1.
 DR Pfam; PF00078; RVT; 1.
 DR KW Hypothetical protein; RNA-directed DNA polymerase; Transferase.
 SQ SEQUENCE 1275 AA; 149009 MW; B327D9D50A581764 CRC64;

Query Match 52.1%; Score 75; DB 2; Length 1275;
 Best Local Similarity 62.5%; Pred. No. 0.016; Mismatches 6; Indels 0; Gaps 0;
 Matches 15; Conservative 3;

QY 2 AQATKATIDKWNCKIKLIFYTSKK 25
 |||||:|||||:|:|:
 Db 1001 AMATKVIDKWDLIKLSFCTAKE 1024

RESULT 14

O00549
 ID O00549 PRELIMINARY; PRT; 573 AA.
 AC O00549;
 DT 01-JUL-1997 (TRENBLrel. 04, Created)
 DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE ORF2-like protein (fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Small-cell lung cancer;
 RX MEDLINE=99119049; PubMed=9918768;
 RA Kuo K.W., Sheu H.M., Huang Y.S., Leung W.C.;
 RT "Expression of transposon LINE-1 is relatively human-specific and
 RT function of the transcripts may be proliferation-essential.";
 RL Biochem. Biophys. Res. Commun. 253:566-570(1998).
 DR EMBL; AF03535; AAD04635.1; -;
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR InterPro; IPR000477; RVTse.

DR Pfam; PF00078; RVT; 1.
 KW RNA-directed DNA polymerase; Transferase.
 FT NON_TER 1
 FT NON_TER 573
 SQ SEQUENCE 573 AA; 66860 MW; A28060BC473AA886 CRC64;

Query Match 50.7%; Score 73; DB 2; Length 573;
 Best Local Similarity 62.5%; Pred. No. 0.015; Mismatches 6; Indels 0; Gaps 0;
 Matches 15; Conservative 3;

QY 2 AQATKATIDKWNCKIKLIFYTSKK 25
 |||||:|||||:|:|:
 Db 515 AMATKIDKWDLIKLSFCTAKE 538

RESULT 15

Q7KZ41
 ID Q7KZ41 PRELIMINARY; PRT; 1192 AA.
 AC Q7KZ41;
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE Hypothetical protein (fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89233117; PubMed=2497061;
 RA Woods-Samuels P., Wong C., Mathias S.L., Scott A.F.,
 RA Kazarian H.H. Jr., Antonarakis S.E.;
 RT "Characterization of a nondeleterious L1 insertion in an intron of the
 RT human factor VIII gene and further evidence of open reading frames in
 RT functional L1 elements.";
 RL Genomics 4:290-296(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Woods-Samuels P.;
 RL Submitted (APR-1989) to the EMBL/GenBank/DBJ databases.
 DR EMBL; M22333; AAA88037.1; -;
 DR InterPro; IPR005135; Exo_endo_phos.
 DR InterPro; IPR000477; RVTse.
 DR Pfam; PF03372; Exo_endo_phos; 1.
 DR Pfam; PF00078; RVT; 1.
 KW Hypothetical protein; RNA-directed DNA polymerase; Transferase.
 FT NON_TER 1
 SQ SEQUENCE 1192 AA; 139677 MW; D706D841DAE0DAD4 CRC64;

Query Match 50.7%; Score 73; DB 2; Length 1192;
 Best Local Similarity 62.5%; Pred. No. 0.03; Mismatches 6; Indels 0; Gaps 0;
 Matches 15; Conservative 3;

QY 2 AQATKATIDKWNCKIKLIFYTSKK 25
 |||||:|||||:|:|:
 Db 918 AMATKIDKWDLIKLSFCTAKE 941

Search completed: November 10, 2004, 14:50:30
 Job time : 38.2981 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 8.05283 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-61

Perfect score: 113

Sequence: 1 VVDVPDFIVWLEAVSDLRAL 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:**
- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
 - 2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
 - 3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
 - 4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
 - 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
 - 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	47	41.6	557	4	US-09-138-277C-3		Sequence 3, Appli
2	46	40.7	126	4	US-09-107-532A-4794		Sequence 4794, Ap
3	46	40.7	304	1	US-08-569-806-6		Sequence 6, Appli
4	45.5	40.3	354	4	US-08-252-991A-27916		Sequence 27916, A
5	45	39.8	365	4	US-09-328-352-6107		Sequence 6107, Ap
6	45	39.8	405	4	US-09-270-767-44068		Sequence 44068, A
7	44	38.9	299	4	US-09-252-991A-18741		Sequence 18741, A
8	44	38.9	336	4	US-09-489-039A-9200		Sequence 9200, Ap
9	44	38.9	351	4	US-08-538-092-1255		Sequence 1255, Ap
10	44	38.9	1003	4	US-09-489-039A-12357		Sequence 12357, A
11	43.5	38.5	691	3	US-09-134-001C-4675		Sequence 4675, Ap
12	43	38.1	171	4	US-09-489-039A-9952		Sequence 9952, Ap
13	43	38.1	375	4	US-09-540-236-2085		Sequence 2085, Ap
14	42.5	37.6	164	4	US-09-107-532A-6548		Sequence 6548, Ap
15	42	37.2	37	4	US-09-716-129-100		Sequence 100, Ap
16	42	37.2	145	4	US-08-716-129-96		Sequence 96, Appli
17	42	37.2	180	4	US-08-543-681A-7612		Sequence 7612, Ap
18	42	37.2	205	4	US-09-489-039A-11450		Sequence 11450, A
19	42	37.2	267	4	US-09-489-039A-12889		Sequence 12889, A
20	42	37.2	282	4	US-08-270-767-41601		Sequence 41601, A
21	42	37.2	481	3	US-09-537-357-7		Sequence 7, Appli
22	42	37.2	487	3	US-09-537-357-15		Sequence 15, Appli
23	42	37.2	943	3	US-09-397-885-5		Sequence 5, Appli
24	42	37.2	943	4	US-09-969-362-5		Sequence 326, App
25	42	37.2	946	4	US-09-492-709A-326		Sequence 141, App
26	41.5	36.7	404	4	US-09-489-847-141		Sequence 141, App
27	41.5	36.7	489	4	US-09-489-847-289		Sequence 289, App

ALIGNMENTS

RESULT 1

US-09-138-277C-3

; Sequence 3, Application US/09138277C

; Patent No. 6436403

; GENERAL INFORMATION:

; APPLICANT: NAKATA, MOTOMI

; APPLICANT: NAKANO, HIROYASU

; APPLICANT: YAGITA, HIDEO

; APPLICANT: OKUMURA, KO

; TITLE OF INVENTION: TRAP FAMILY MOLECULES, POLYNUCLEOTIDES ENCODING THEM,

; TITLE OF INVENTION: AND ANTIBODIES AGAINST THEM

; FILE REFERENCE: 007898-0255515

; CURRENT APPLICATION NUMBER: US/09138,277C

; CURRENT FILING DATE: 1998-08-18

; PRIOR APPLICATION NUMBER: PCT/JP97/00512

; PRIOR FILING DATE: 1997-02-24

; PRIOR APPLICATION NUMBER: JP 34674/1996

; PRIOR FILING DATE: 1996-02-22

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 3

; LENGTH: 557

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-138-277C-3

Query Match 41.6%; Score 47; DB 4; Length 557;

Best Local Similarity 64.3%; Pred. No. 38;

Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 VMLEAVSDLRAL 22

Db 258 VQLEQISDLRSL 271

RESULT 2

US-09-107-532A-4794

; Sequence 4794, Application US/09107532A

; Patent No. 6583275

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

; NUMBER OF SEQUENCES: 7310

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA


```
Query Match 39.8%; Score 45; DB 4; Length 365;
Best Local Similarity 61.5%; Pred. No. 48;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 DFIVLEEAASDL 18
   |||:||||:|
Db 199 DFLVLEENMDGL 211

RESULT 6
US-09-270-767-44068
; Sequence 44068, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44068
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-44068

Query Match 39.8%; Score 45; DB 4; Length 405;
Best Local Similarity 57.1%; Pred. No. 54;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 IWLLEEAASDLHRA 21
   |||:||||:|
Db 383 IWLIAECVAQKRA 396

RESULT 7
US-09-252-991A-18741
; Sequence 18741, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18741
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18741

Query Match 38.9%; Score 44; DB 4; Length 299;
Best Local Similarity 75.0%; Pred. No. 55;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 11 LEEAVSGLHRL 22
   |||:||||:|
Db 76 VEEAVAGLHRL 87

RESULT 8
US-09-489-039A-9200
; Sequence 9200, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
```

```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9200
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9200

Query Match 38.9%; Score 44; DB 4; Length 336;
Best Local Similarity 47.1%; Pred. No. 63;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 VVDVDFIVWLEEAASD 17
   |||:||||:|
Db 30 VMKVSDDLTLAEAILD 46

RESULT 9
US-09-538-092-1256
; Sequence 1256, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqformatter Version 0.9
; SEQ ID NO 1256
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q02547
US-09-538-092-1256

Query Match 38.9%; Score 44; DB 4; Length 351;
Best Local Similarity 50.0%; Pred. No. 66;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 8 IWLLEEAASDLHRA 21
   |||:||||:|
Db 329 IWLIAECIAQRHRA 342

RESULT 10
US-09-489-039A-12357
; Sequence 12357, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
```

; SEQ ID NO 12357
; LENGTH: 1003
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12357

Query Match 38.9%; Score 44; DB 4; Length 1003;
Best Local Similarity 45.5%; Pred. No. 2.1e+02;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 VDPVDFVWLEAEVSDLHRL 22
DB 313 VVDEAHLVWSEAFSREIQAI 334

RESULT 11

US-09-134-001C-4675
; Sequence 4675, Application US/09134001C
; Patent No. 6380370

GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4675
; LENGTH: 691
; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4675

Query Match 38.5%; Score 43.5; DB 3; Length 691;
Best Local Similarity 45.0%; Pred. No. 1.7e+02;
Matches 9; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 4 VPDFIV-WLEAEVSDLHRL 22
DB 357 IPDFVLEQLFEHINDLHKDL 376

RESULT 12

US-09-489-039A-9952
; Sequence 9952, Application US/09489039A
; Patent No. 6610836

GENERAL INFORMATION:

; APPLICANT: Gary Bretton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9952
; LENGTH: 171
; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9952

Query Match 38.1%; Score 43; DB 4; Length 171;
Best Local Similarity 44.4%; Pred. No. 43;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 DVPDFVWLEAEVSDLHR 20
DB 67 DLPDSIVPLVYEGIKDMYQ 104

RESULT 13

US-09-540-236-2085
; Sequence 2085, Application US/09540236
; Patent No. 6673910

GENERAL INFORMATION:

; APPLICANT: Gary L. Bretton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2085
; LENGTH: 375
; TYPE: PRT

; ORGANISM: M.catarrhalis
US-09-540-236-2085

Query Match 38.1%; Score 43; DB 4; Length 375;
Best Local Similarity 42.9%; Pred. No. 1e+02;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 6 DFVWLEAEVSDLH 19
DB 202 DFVWLEAEVSDLH 215

RESULT 14

US-09-107-532A-6548
; Sequence 6548, Application US/09107532A
; Patent No. 6583275

GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSER: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 6548:
SEQUENCE CHARACTERISTICS:
LENGTH: 164 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:

```

; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...164
; SEQUENCE DESCRIPTION: SEQ ID NO: 6548:
US-09-107-532A-6548

```

```

Query Match      37.6%; Score 42.5; DB 4; Length 164;
Best Local Similarity 27.6%; Pred. No. 49;
Matches 8; Conservative 8; Mismatches 6; Indels 7; Gaps 1;

```

```

QY 1 VVDVPDFI-----VMLEEAVSDLHRL 22
Db 78 VVEIPAVVGATGVEPITLRKFISDFHKG 106

```

```

RESULT 15
US-09-716-129-100
; Sequence 100, Application US/09716129
; Patent No. 6632920
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2025P1
; CURRENT APPLICATION NUMBER: US/09/716,129
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/076,053
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,057
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,052
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,054
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,051
; PRIOR FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 100
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-716-129-100

```

```

Query Match      37.2%; Score 42; DB 4; Length 37;
Best Local Similarity 36.8%; Pred. No. 11;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 4 VPDFIVMLEEAVSDLHRL 22
Db 9 LPDYLAIDEALAHVTL 27

```

```

Search completed: November 10, 2004, 14:55:43
Job time : 8.10283 secs

```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 25.4453 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-61
Perfect score: 113
Sequence: 1 VVDVPDFVWLEAVSDLHRAL 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113	100.0	22	14	US-10-092-750-61
2	61	54.0	22	14	US-10-092-750-227
3	50	44.2	351	14	US-10-369-493-11619
4	50	44.2	355	14	US-10-369-493-14336
5	50	44.2	355	14	US-10-369-493-14303
6	50	44.2	355	14	US-10-369-493-14999
7	49	43.4	790	16	US-10-437-963-156459
8	48	42.5	611	15	US-10-382-122A-69638
9	47	41.6	78	16	US-10-437-963-124299
10	47	41.6	451	14	US-10-004-378A-2
11	47	41.6	538	15	US-10-042-865-163
12	47	41.6	557	14	US-10-004-378A-34
13	47	41.6	557	15	US-10-042-865-162

14	47	41.6	568	15	US-10-262-445-132	Sequence 132, Appl
15	47	41.6	568	15	US-10-042-865-38	Sequence 38, Appl
16	47	41.6	1022	14	US-10-156-761-9292	Sequence 9292, Ap
17	46	40.7	848	9	US-09-882-986-2	Sequence 2, Appl
18	46	40.7	848	15	US-10-469-013-18	Sequence 18, Appl
19	46	40.7	1295	16	US-10-437-963-144134	Sequence 144134,
20	46	40.7	3562	14	US-10-341-434-109	Sequence 109, App
21	45.5	40.3	327	17	US-10-435-115-199784	Sequence 195784,
22	45.5	40.3	541	16	US-10-437-963-151993	Sequence 151993,
23	45	39.8	240	15	US-10-282-122A-47961	Sequence 47961, A
24	45	39.8	266	14	US-10-369-493-8199	Sequence 8199, Ap
25	45	39.8	402	14	US-10-369-493-18774	Sequence 18774, A
26	45	39.8	1004	14	US-10-156-761-14806	Sequence 14806, A
27	44.5	39.4	317	14	US-10-369-493-1258	Sequence 1258, Ap
28	44.5	39.4	317	14	US-10-369-493-20325	Sequence 20325, A
29	44.5	39.4	317	14	US-10-369-493-21652	Sequence 21652, A
30	44.5	39.4	543	14	US-10-369-493-17936	Sequence 17936, A
31	44	38.9	118	15	US-10-424-599-175022	Sequence 175022,
32	44	38.9	129	16	US-10-437-963-165340	Sequence 165340,
33	44	38.9	175	15	US-10-424-599-253697	Sequence 253697,
34	44	38.9	232	15	US-10-276-774-2499	Sequence 2499, Ap
35	44	38.9	234	17	US-10-489-372-23	Sequence 23, Appl
36	44	38.9	345	14	US-10-365-646-4	Sequence 4, Appl
37	44	38.9	345	14	US-10-365-646-5	Sequence 5, Appl
38	44	38.9	351	10	US-09-997-003-33	Sequence 33, Appl
39	44	38.9	400	14	US-10-369-493-20208	Sequence 20208, A
40	44	38.9	461	14	US-10-176-306-23	Sequence 23, Appl
41	44	38.9	461	14	US-10-094-749-2145	Sequence 2145, Ap
42	44	38.9	461	14	US-10-285-027-305	Sequence 205, App
43	44	38.9	461	15	US-10-188-832-97	Sequence 97, Appl
44	44	38.9	906	16	US-10-437-963-105322	Sequence 105322,
45	44	38.9	943	17	US-10-425-115-278341	Sequence 278341,

ALIGNMENTS

RESULT 1
US-10-092-750-61
; Sequence 61, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-61

Query Match 100.0%; Score 113; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 8e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VVDVPDFVWLEAVSDLHRAL 22
Db 1 VVDVPDFVWLEAVSDLHRAL 22
RESULT 2
US-10-092-750-227
; Sequence 227, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:

```

; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 227
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-227

Query Match          54.0%; Score 61; DB 14; Length 22;
Best Local Similarity 75.0%; Pred. No. 0.048;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 DVDPDFIVWLEAVSDLHRAL 22
DB 3 DVVGFIDSELEAVSDLHRAL 22

RESULT 3
US-10-369-493-11619
; Sequence 11619, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11619
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-11619

Query Match          44.2%; Score 50; DB 14; Length 351;
Best Local Similarity 44.4%; Pred. No. 40;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 VVDVPDFIVWLEAVSDL 18
DB 190 LIDKPDFFFWLEKNWDDI 207

RESULT 4
US-10-369-493-14336
; Sequence 14336, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14999
; LENGTH: 355
; TYPE: PRT

; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14336
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14336

Query Match          44.2%; Score 50; DB 14; Length 355;
Best Local Similarity 44.4%; Pred. No. 41;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 VVDVPDFIVWLEAVSDL 18
DB 188 LIDKPDFFFWLEKNWDDI 205

RESULT 5
US-10-369-493-14603
; Sequence 14603, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14603
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14603

Query Match          44.2%; Score 50; DB 14; Length 355;
Best Local Similarity 44.4%; Pred. No. 41;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 VVDVPDFIVWLEAVSDL 18
DB 188 LIDKPDFFFWLEKNWDDI 205

RESULT 6
US-10-369-493-14999
; Sequence 14999, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14999
; LENGTH: 355
; TYPE: PRT

```

```
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14999

Query Match          44.2%; Score 50; DB 14; Length 355;
Best Local Similarity 44.4%; Pred. No. 41;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 VVDVDFIVWLEEA VSDL 18
   :|||||:|:|:|:|:|
Db 188 LIDKDFFFWLEK WDDI 205

RESULT 7
US-10-437-963-156459
; Sequence 156459, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 156459
; LENGTH: 790
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_56124C.1.pep
US-10-437-963-156459

Query Match          43.4%; Score 49; DB 16; Length 790;
Best Local Similarity 53.3%; Pred. No. 1.4e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 DVPDFIVWLEEA VSD 17
   :|||||:|:|:|:|
Db 627 DQPDFISWVSECLD 641

RESULT 8
US-10-282-122A-69638
; Sequence 69638, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
```

```
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69638
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-10-282-122A-69638

Query Match          42.5%; Score 48; DB 15; Length 611;
Best Local Similarity 52.6%; Pred. No. 1.5e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 DVPDFIVWLEEA VSDLHRA 21
   :|||||:|:|:|:|
Db 139 DTFDLAVLKSAVKELHGA 157

RESULT 9
US-10-437-963-124299
; Sequence 124299, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 124299
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_27051C.1.pep
US-10-437-963-124299

Query Match          41.6%; Score 47; DB 16; Length 78;
Best Local Similarity 69.2%; Pred. No. 23;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 WLEEA VSDLHRA 21
   :|||||:|:|:|:|
Db 41 LWLEA FSPFLHRA 53

RESULT 10
US-10-004-378A-2
; Sequence 2, Application US/10004378A
```

```

; Publication No. US20030228301A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Furtak, Kazaryna
; APPLICANT: Perna, Amanda
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Casman, Stacie J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Vernet, Corinne A
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Agee, Michele
; APPLICANT: Rastelli, Luca
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Gerlach, Valerie
; APPLICANT: Edinger, Schlomit
; APPLICANT: MacDougall, John R
; APPLICANT: Peyman, John A
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David J
; APPLICANT: Ellerman, Karen
; APPLICANT: Gangolli, Esha A
; TITLE OF INVENTION: No. US20030228301A1el Human Proteins, Polynucleotides Encoding TH
; FILE REFERENCE: 21402-179
; CURRENT APPLICATION NUMBER: US/10/004,378A
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 60/242,882
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,765
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/300,206
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/242,789
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,768
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,767
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/243,622
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/273,047
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/243,591
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,950
; PRIOR FILING DATE: 2000-10-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-004-378A-2

Query Match 41.6%; Score 47; DB 14; Length 451;
Best Local Similarity 54.3%; Pred. No. 1.5e+02;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 VMLEEAVSLHRL 22
DB 152 VQLEEQISDLHKS 165

RESULT 11
US-10-042-865-163

```

```

; Sequence 163, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corinne A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Ference L
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Schlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 163
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-865-163

Query Match 41.6%; Score 47; DB 15; Length 538;
Best Local Similarity 64.3%; Pred. No. 1.8e+02;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 VMLEEAVSLHRL 22
DB 239 VQLEEQISDLHKS 252

RESULT 12
US-10-004-378A-34
; Sequence 34, Application US/10004378A
; Publication No. US20030228301A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Furtak, Kazaryna
; APPLICANT: Perna, Amanda
; APPLICANT: Patturajan, Meera

```



```

; APPLICANT: Shimkets, Richard A
; APPLICANT: Guo, Xiaojia Sashia
; APPLICANT: Casman, Stacie J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Vernet, Corrine A
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Agee, Michele
; APPLICANT: Rastelli, Luca
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Gerlach, Valerie
; APPLICANT: Edinger, Schlomit
; APPLICANT: MacDougall, John R
; APPLICANT: Peyman, John A
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David J
; APPLICANT: Ellerman, Karen
; APPLICANT: Gangolli, Esha A
; TITLE OF INVENTION: No. US2003028301A1el Human Proteins, Polynucleotides Encoding Them
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 21402-179
; CURRENT APPLICATION NUMBER: US/10/004,378A
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 60/242,882
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,765
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/300,206
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/242,789
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,768
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,767
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/243,622
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: 60/273,047
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/243,591
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,950
; PRIOR FILING DATE: 2000-10-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-004-378A-34

Query Match 41.6%; Score 47; DB 14; Length 557;
Best Local Similarity 64.3%; Pred. No. 1.9e+02;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 VMLEEAVSDLHRL 22
DB 258 VQLEEQISDLHKS 271

RESULT 13
US-10-042-865-162
; Sequence 162, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D

; APPLICANT: Shimkets, Richard A
; APPLICANT: Guo, Xiaojia Sashia
; APPLICANT: Casman, Stacie J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Vernet, Corrine A
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Agee, Michele
; APPLICANT: Rastelli, Luca
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Gerlach, Valerie
; APPLICANT: Edinger, Schlomit
; APPLICANT: MacDougall, John R
; APPLICANT: Peyman, John A
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David J
; APPLICANT: Ellerman, Karen
; APPLICANT: Gangolli, Esha A
; TITLE OF INVENTION: No. US2003028301A1el Human Proteins, Polynucleotides Encoding Them and Methods of
; TITLE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-337
; CURRENT APPLICATION NUMBER: US/10/042,865
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 162
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-865-162

Query Match 41.6%; Score 47; DB 15; Length 557;
Best Local Similarity 64.3%; Pred. No. 1.9e+02;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 VMLEEAVSDLHRL 22
DB 258 VQLEEQISDLHKS 271

RESULT 14
US-10-262-445-132
; Sequence 132, Application US/10262445
; Publication No. US20040014058A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John
; APPLICANT: Burgess, Catherine
; APPLICANT: Catterton, Elina
; APPLICANT: Chant, John
; APPLICANT: Chaudhuri, Amitabha
; APPLICANT: Edinger, Schlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Giot, Loic
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia

```

```

; APPLICANT: Kekuda, Ramesh
; APPLICANT: Mezes, Peter
; APPLICANT: Millet, Isabelle
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Patturajan, Meera
; APPLICANT: Rieger, Daniel
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Zhong, Hailong
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 21402-462D
; CURRENT APPLICATION NUMBER: US/10/262,445
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/327,454
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/329,414
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/330,142
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 60/341,058
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/343,629
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 60/349,575
; PRIOR FILING DATE: 2001-10-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 132
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-262-445-132

Query Match 41.6%; Score 47; DB 15; Length 568;
Best Local Similarity 64.3%; Pred. No. 1.9e+02;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 VWLEAVSDLHRL 22
DB 269 VQLEEQISDLHKS 282

RESULT 15
US-10-042-865-38
; Sequence 38, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchenev, Velizar T
; APPLICANT: Miller, Charles E

```

```

; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Perence L
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Mayankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; TITLE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 38
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-042-865-38

Query Match 41.6%; Score 47; DB 15; Length 568;
Best Local Similarity 64.3%; Pred. No. 1.9e+02;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 VWLEAVSDLHRL 22
DB 269 VQLEEQISDLHKS 282

```

Search completed: November 11, 2004, 02:43:17
Job time : 26.4953 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 / Search time 5.39623 Seconds
(without alignments)

392.268 Million cell updates/sec

Title: US-10-092-750-61

Perfect score: 113

Sequence: 1 VVDVPDFIVWLEFAVSDLRAL 22

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54.5	48.2	349	2 T25348	hypothetical prote
2	53	46.9	806	2 H70647	probable NADH2 de
3	51	45.1	349	2 T26367	hypothetical prote
4	50	44.2	294	2 AG0913	probable adenine-s
5	50	44.2	363	2 AG3002	3-dehydroquinat
6	50	44.2	377	2 AG9281	3-dehydroquinat
7	48.5	42.9	328	2 F33599	hypothetical prote
8	48	42.5	82	2 AB0272	hypothetical prote
9	47	41.6	292	2 A84487	modification methy
10	47	41.6	557	2 JC6539	tumor necrosis fac
11	46	40.7	197	2 A87286	conserved hypotet
12	46	40.7	296	2 F91145	probable methyltra
13	46	40.7	296	2 B85991	hypothetical adeni
14	46	40.7	296	2 H55118	3-dehydroquinat
15	46	40.7	378	2 A32527	probable cap-bindi
16	46	40.7	749	2 C84508	hypothetical prote
17	46	40.7	829	2 T45036	dystrophin - mouse
18	46	40.7	3678	2 S28916	dystrophin, muscle
19	46	40.7	3685	1 A37605	hypothetical prote
20	45	39.8	163	2 B84308	hypothetical prote
21	45	39.8	536	2 T41652	carbon dioxide con
22	45	39.8	539	2 D36904	hypothetical prote
23	44.5	39.4	109	2 S38836	ornithine carbamoy
24	44.5	39.4	317	2 G75041	probable ornithine
25	44.5	39.4	317	2 G75041	hypothetical prote
26	44	38.9	119	2 T22181	probable dihydrone
27	44	38.9	124	2 B71494	hypothetical prote
28	44	38.9	129	2 F81173	hypothetical prote
29	44	38.9	129	2 DB1931	hypothetical prote

30 44 38.9 139 2 E81173
31 44 38.9 202 2 AF2831
32 44 38.9 202 2 B97609
33 44 38.9 207 2 C69338
34 44 38.9 271 2 C47127
35 44 38.9 273 2 A32123
36 44 38.9 274 2 JN0908
37 44 38.9 386 2 AE1853
38 44 38.9 441 2 T49113
39 44 38.9 493 2 T22180
40 44 38.9 563 2 F87986
41 44 38.9 1140 2 T41457
42 44 38.9 1203 2 T04294
43 44 38.9 1489 2 G71406
44 44 38.9 1633 2 T01879
45 43.5 38.5 467 1 C69301

ALIGNMENTS

RESULT 1

T25348

hypothetical protein T27C5.5 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T25348

R;Cummings, P.

submitted to the EMBL Data Library, November 1996

A;Reference number: Z20020

A;Accession: T25348

A;Status: preliminary; translated from GB/EMBL/DBDJB

A;Molecule type: DNA

A;Residues: 1-349 <WIL>

A;Cross-references: UNIPROT:O45849; EMBL:Z82058; PIDN:CAB04866.1; GSPDB:GN00023; CESP:T

A;Experimental source: clone T27C5

C;Genetics:

A;Gene: CESP.T27C5.5

A;Map position: 5

A;Introns: 58/2; 171/3; 273/1

C;Superfamily: Caenorhabditis hypothetical protein C49G7.2

Query Match 48.2%; Score 54.5; DB 2; Length 349;

Best Local Similarity 43.5%; Pred. No. 2.3;

Matches 10; Conservative 6; Mismatches 6; Indels 1; Gaps 1;

Qy 1 VVDVPDFIVWLE-AVSDLRAL 22

Db 180 IIDDPDFHWEPEPMYAIHYAI 202

RESULT 2

H70647

probable NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain G - Mycobacterium tuberc

C;Species: Mycobacterium tuberculosis

C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C;Accession: H70647

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrrell, B.G.

A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A;Reference number: A70500; MUID:98295987; PMID:9834230

A;Accession: H70647

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-806 <COL>

A;Cross-references: UNIPROT:P95175; GB:Z83867; GB:AL123456; NID:G3261695; PIDN:CAB06281

A;Experimental source: strain H37Rv

C;Genetics:

A;Gene: nuoG

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 11

A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
A;Reference number: A62950; MUID:20437337; PMID:10984043
A;Accession: F83599
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-328 <STO>
A;Cross-references: UNIPROT:Q916D4; GB:AE004474; GB:AE004091; NID:g9946210; PIDN:AA0374
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA0360

Query Match 42.9%; Score 48.5; DB 2; Length 328;
Best Local Similarity 64.7%; Pred. No. 17;
Matches 11; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

Qy 4 VPDFVWLEAVSLHR 20
||| | | | | | | | |
Db 15 VPSP-AWAEDAVSLAR 30

RESULT 8
A60272
hypothetical protein YPO2231 [imported] - *Yersinia pestis* (strain CO92)
C;Species: *Yersinia pestis*
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: A60272
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, P.; Dougan, G.; Li, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001
A;Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: A60272
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-82 <KUR>
A;Cross-references: UNIPROT:Q8ZEE2; GB:AL590842; PIDN:CAC91038.1; PID:g15980232; GSPDB:G
C;Genetics:
A;Gene: YPO2231

Query Match 42.5%; Score 48; DB 2; Length 82;
Best Local Similarity 47.1%; Pred. No. 4.6;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VVDVDFVWLEAVSL 17
||| | | | | | | | |
Db 66 MLDVTDLSLWLEKPVQD 82

RESULT 9
A64487
modification methylase (EC 2.1.1.1-) homolog - *Methanococcus jannaschii*
C;Species: *Methanococcus jannaschii*
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: A64487
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.J.; Overbeek, R.; Kirkness, E.P.; Weisscock, K.G.; Merrick, J.M.; Glodek, A.; rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A. Science 273, 1058-1073, 1996
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C. A;Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*
A;Reference number: A64300; MUID:96337999; PMID:8688087
A;Accession: A64487
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-292 <BUL>
A;Cross-references: UNIPROT:Q58893; GB:U67590; GB:L77117; NID:g1592126; PIDN:AAB99509.1.
C;Genetics:
A;Map position: F081470773-1471651
C;Superfamily: type II site-specific DNA-methyltransferase
C;Keywords: methyltransferase

Query Match 41.6%; Score 47; DB 2; Length 292;

Best Local Similarity 41.2%; Pred. No. 24;
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
Qy 6 DFVWLEAVSLHRAL 22
||| | | | | | | | |
Db 49 DYLNWIEVVVKIKVVL 65

RESULT 10
JC6539
tumor necrosis factor receptor-associated factor 5 homolog - human
C;Species: *Homo sapiens* (man)
C;Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 09-Jul-2004
C;Accession: JC6539
R;Mizushima, S.; Fujita, M.; Ishida, T.; Azuma, S.; Kato, K.; Hirai, M.; Otsuka, M.; Yano, A.; Title: Cloning and characterization of a cDNA encoding the human homolog of tumor necrosis factor receptor-associated factor 5
A;Reference number: JC6539; MUID:98172745; PMID:9511754
A;Accession: JC6539
A;Molecule type: mRNA
A;Residues: 1-557 <MIZ>
A;Cross-references: UNIPROT:Q00463; DDBJ:AB000509; NID:g2982670; PIDN:BAA25262.1; PID:g
C;Comment: This protein is involved in transduction of signals from various tumor necrosis factor receptor-beta receptor.
C;Genetics:
A;Map position: 1q32.3-1q41.1
C;Superfamily: TNF receptor-associated factor (TRAF); RING finger homology
C;Keywords: coiled coil; tumor; zinc finger
F;41-90/Domain: RING finger homology <RRN>

Query Match 41.6%; Score 47; DB 2; Length 557;
Best Local Similarity 64.3%; Pred. No. 48;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 9 VWLEAVSLHRAL 22
||| | | | | | | | |
Db 258 VQLEQISDLHKSL 271

RESULT 11
A87286
conserved hypothetical protein CC0298 [imported] - *Caulobacter crescentus*
C;Species: *Caulobacter crescentus*
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: A87286
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolc n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of *Caulobacter crescentus*.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: A87286
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-197 <STO>
A;Cross-references: UNIPROT:Q9ABD1; GB:AE005673; NID:g13421439; PIDN:AAK22285.1; GSPDB:G
C;Genetics:
A;Gene: CC0298

Query Match 40.7%; Score 46; DB 2; Length 197;
Best Local Similarity 81.8%; Pred. No. 23;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 12 EEAVSLHRAL 22
||| | | | | | | | |
Db 152 EEAVSLHRRM 162

RESULT 12
F91145
probable methyltransferase [imported] - *Escherichia coli* (strain O157:H7, substrain R1N
C;Species: *Escherichia coli*
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C:Accession: F91145
 R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic islands
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: F91145
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-236 <HAY>
 A:Cross-references: UNIPROT:Q8X9A6; GB:BA000007; PIDN:BA37557.1; PID:gl13363607; GSPDB:G
 A:Experimental source: strain O157:H7, substrain RMD 050952
 C:Genetics:
 A:Gene: ECs4134
 C:Superfamily: type II site-specific DNA-methyltransferase

Query Match 40.7%; Score 46; DB 2; Length 296;
 Best Local Similarity 50.0%; Pred. No. 35;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 7 FIVWLEAVSDHLRAL 22
 ||||| : : : :
 DB 63 FIDWLLEVIAECHRVL 78

RESULT 13

B85991
 probable methyltransferase yhdJ [imported] - *Escherichia coli* (strain O157:H7, substrain
 C:Species: *Escherichia coli*
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C:Accession: B85991
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: B85991
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-296 <STO>
 A:Cross-references: UNIPROT:Q8X9A6; GB:AE005174; NID:gl2517892; PIDN:AAG58390.1; GSPDB:G
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: yhdJ
 C:Superfamily: type II site-specific DNA-methyltransferase

Query Match 40.7%; Score 46; DB 2; Length 296;
 Best Local Similarity 50.0%; Pred. No. 35;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 7 FIVWLEAVSDHLRAL 22
 ||||| : : : :
 DB 63 FIDWLLEVIAECHRVL 78

RESULT 14

H65118
 hypothetical adenine-specific methylase in *fts*-envR intergenic region - *Escherichia coli*
 C:Species: *Escherichia coli*
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
 A:Accession: H65118; D47043; C28207
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of *Escherichia coli* K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: H65118
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-236 <BLAT>
 A:Cross-references: UNIPROT:P28638; GB:AE000405; GB:U00096; NID:gl1789659; PIDN:AAC76294.
 A:Experimental source: strain K-12, substrain MG1655
 R:Ball, C.A.; Osuna, R.; Ferguson, K.C.; Johnson, R.C.

J. Bacteriol. 174, 8043-8056, 1992
 A:Title: Dramatic changes in *Fis* levels upon nutrient upshift in *Escherichia coli*.
 A:Reference number: A47043; MUID:93094136; PMID:1459953
 A:Accession: D47043
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 3-29, 'N', 31-102 <BAL>
 A:Cross-references: GB:M95784; NID:gl45974; PIDN:AAA23784.1; PID:gl145978
 A:Note: Sequence extracted from NCBI backbone (NCBIN:119972, NCBIP:119976)
 R:Johnson, R.C.; Ball, C.A.; Pfeiffer, D.; Simon, M.I.
 Proc. Natl. Acad. Sci. U.S.A. 85, 3484-3488, 1988
 A:Title: Isolation of the gene encoding the *Hlr* recombinational enhancer binding protein
 A:Reference number: A28207; MUID:88217925; PMID:2835774
 A:Accession: C28207
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 3-13 <JOH>
 A:Cross-references: GB:J03245; NID:gl45970; PIDN:AAA93857.1; PID:gl128954
 C:Genetics:
 A:Gene: yhdJ
 C:Superfamily: type II site-specific DNA-methyltransferase

Query Match 40.7%; Score 46; DB 2; Length 296;
 Best Local Similarity 50.0%; Pred. No. 35;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 7 FIVWLEAVSDHLRAL 22
 ||||| : : : :
 DB 63 FIDWLLEVIAECHRVL 78

RESULT 15

AF3257
 3-dehydroquininate synthase (EC 4.2.3.4) [imported] - *Brucella melitensis* (strain 16M)
 C:Species: *Brucella melitensis*
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
 C:Accession: AF3257
 R:DeIvecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Muijer, C.; Los, T.; Ivanova,
 ; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen *Brucella meliten*
 A:Reference number: AD3252; PMID:11756688
 A:Accession: AF3257
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-378 <KUR>
 A:Cross-references: UNIPROT:Q8YJN9; GB:AE008917; PIDN:AAL51225.1; PID:gl17981913; GSPDB:
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BMEI0043
 A:Map position: 1

Query Match 40.7%; Score 46; DB 2; Length 378;
 Best Local Similarity 58.3%; Pred. No. 45;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VDVDPDFIWLLE 12
 : : : : : :
 DB 200 LIDRDPDFFAWLE 211

Search completed: November 10, 2004, 14:52:28
 Job time : 6.39623 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 29.3057 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-61
Perfect score: 113
Sequence: 1 VVDVDPFIVMLEEAVSDDLHRAL 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113	100.0	22	Q96JA7	Q96ja7 homo sapien
2	54.5	48.2	349	Q45849	Q45849 caenorhabdi
3	54	47.8	357	Q7Y3V5	Q7y3v5 bacterioph
4	53	46.9	806	1 NUOG MYCBO	P59662 mycobacteri
5	53	46.9	806	1 NUOG MYCTU	P95175 mycobacteri
6	52	46.0	3094	2 Q8UDR3	Q8udr3 plasmodium
7	52	46.0	3094	2 AAQ73930	AAq73930 plasmodiu
8	51	45.1	349	2 Q9XX66	Q9xx66 caenorhabdi
9	51	45.1	894	2 Q7PZJ5	Q7pzj5 anopheles g
10	50	44.2	227	2 Q6RFG9	Q6rfg9 pycnococcus
11	50	44.2	227	2 AAR99499	Aar99499 pycnococc
12	50	44.2	294	2 Q8XF74	Q8xf74 salmonella
13	50	44.2	294	2 Q7CPL9	Q7cpl9 salmonella
14	50	44.2	377	1 AROB-AGRT5	Q8u9v0 agrobacteri
15	50	44.2	395	2 Q728W5	Q728w5 desulfovibr
16	50	44.2	395	2 AAS96959	Aas96959 desulfovi
17	49	43.4	389	2 Q9LW88	Q9lwe8 oryza sativ
18	49	43.4	3680	1 DMD CANFA	Q97592 canis famil
19	48.5	42.9	328	2 Q91604	Q9i6d4 pseudomonas
20	48.5	42.9	513	2 Q9W067	Q9w067 drosophila
21	48.5	42.9	588	2 Q86IH2	Q86ih2 dictyosteli
22	48	42.5	82	2 Q8ZE32	Q8ze2 dictyosteli
23	48	42.5	82	2 AAS62246	Aas62246 yersinia
24	48	42.5	345	2 Q91UW7	Q91uw7 uncultured
25	48	42.5	610	1 GLMS_PBSBM	Q87tt8 p glucosami
26	48	42.5	856	2 Q7S1G9	Q7s1g9 neurospora
27	48	42.5	1633	2 Q8UDW9	Q8udw9 plasmodium
28	48	42.5	1633	2 AAQ73924	AAq73924 plasmodiu
29	48	42.5	3056	2 Q8I639	Q8i639 plasmodium
30	48	42.5	3064	2 Q6UDW7	Q6udw7 plasmodium
31	48	42.5	3064	2 AAQ73926	AAq73926 plasmodiu

32	47.5	42.0	485	2	Q7Q6Q6	Q7q6q6 anopheles g
33	47	41.6	292	1	MTMS_MFTJA	Q58993 methanococc
34	47	41.6	429	2	Q6PB16	Q6pb16 xenopus lae
35	47	41.6	429	2	AAH59968	AAh59968 xenopus l
36	47	41.6	557	1	TRA5_HUMAN	O00463 homo sapien
37	47	41.6	1022	2	Q82MA7	Q82ma7 streptomyc
38	47	41.6	1234	2	Q8VD34	Q8vd34 mus musculu
39	46.5	41.2	2766	1	THYG_MOUSE	Q08710 mus musculu
40	46	40.7	143	2	Q688N6	Q688n6 plasmodium
41	46	40.7	143	2	AAH32054	AAh32054 plasmodiu
42	46	40.7	166	2	Q89XN3	Q89xn3 bradyrhizob
43	46	40.7	197	2	Q9ABD1	Q9abd1 caulobacter
44	46	40.7	238	2	Q61R98	Q6i98 xenopus lae
45	46	40.7	238	2	AAH71002	AAh71002 xenopus l

ALIGNMENTS

RESULT 1

Q96JA7	PRELIMINARY;	PRT;	22 AA.
ID	Q96JA7		
AC	Q96JA7;		
DT	01-DEC-2001 (Tremblrel. 19, Created)		
DT	01-DEC-2001 (Tremblrel. 19, Last sequence update)		
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)		
DE	Bcl-XL-binding protein v18 (Fragment).		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Theria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Kidney;		
RA	MEDLINE=21293069; PubMed=11283018;		
RA	Hammond P.W.; Alpin J.; Rise C.E.; Wright M.; Kreider B.L.;		
RT	"In Vitro Selection and Characterization of Bcl-XL-binding Proteins		
RT	from a Mix of Tissue-specific mRNA Display Libraries.";		
RL	J. Biol. Chem. 276:20898-20906(2001).		
DR	EMBL; AF357526; AAK60630.1; -.		
FT	NON_TER	1	
FT	NON_TER	22	
SQ	SEQUENCE	22 AA;	2523 MW; 0F87448F26BCF1F9 CRC64;

Query Match 100.0%; Score 113; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.9e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	VVDVDPFIVMLEEAVSDDLHRAL 22
Db	1	VVDVDPFIVMLEEAVSDDLHRAL 22

RESULT 2

C45849	PRELIMINARY;	PRT;	349 AA.
ID	C45849		
AC	C45849;		
DT	01-JUN-1998 (Tremblrel. 06, Created)		
DT	01-JUN-1998 (Tremblrel. 06, Last sequence update)		
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)		
DE	C. elegans SRH-132 protein. (Corresponding sequence T27C5.5).		
GN	Names=srh-132; ORFNames=127C5.5;		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;		
OC	Rhabditidae; Pelodierinae; Caenorhabditis.		
OX	NCBI_TaxID=6239;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Bristol N2;		
RA	MEDLINE=99069613; PubMed=9851916;		
RT	Genome sequence of the nematode C.elegans: A platform for		
RT	investigating biology.";		

June 15

RL	Science 282:2012-2018 (1998).	OS	Mycobacterium bovis.
RN	[2]	OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
RP	SEQUENCE FROM N.A.	OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RC	STRAIN-Bristol N2.	OX	NCBI_TaxID=1765;
RA	Cummings PN.	RN	[1]
RL	Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.	RP	SEQUENCE FROM N.A.
DR	EMBL; 282058; CAS04866.1; -	RC	STRAIN=AF2122/97;
DR	PIR; T25348; T25348.	RX	MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
DR	WormPep; T27C5.5; CE16499.	RA	Garnier T., Eglmeier K., Camus J.-C., Medina N., Mansoor H.,
DR	GO; GO:0016020; C:membrane; IEA.	RA	Pyrro M., Duthoy S., Grondin S., Lacroix C., Mensemp C., Simon S.,
DR	GO; GO:0004930; F:G-protein coupled receptor activity; IEA.	RA	Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
DR	InterPro; IPR003003; 7TM chemrecept2.	RA	Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
DR	InterPro; IPR000168; Nm7TM_chemrecept.	RT	"The complete genome sequence of Mycobacterium bovis.";
DR	Pfam; PF01604; 7tm_5; 1	RL	Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
SQ	SEQUENCE 349 AA; 40102 MW; DB40D702EC098901 CRC64;	CC	-1- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
	Query Match 48.2%; Score 54.5; DB 21; Length 349;	CC	sulfur (Fe-S) centers, to quinones in the respiratory chain. The
	Best Local Similarity 43.5%; Pred. No. 7.9;	CC	immediate electron acceptor for the enzyme in this species is
	Matches 10; Conservative 6; Mismatches 6; Indels 1; Gaps 1;	CC	believed to be menaquinone. Couples the redox reaction to proton
		CC	translocation (for every two electrons transferred, four hydrogen
		CC	ions are translocated across the cytoplasmic membrane), and thus
		CC	conserves the redox energy in a proton gradient (By similarity).
QY	1 VVDVDFIVWLEE-AVSDLHRL 22	CC	-1- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
DB	180 IIDDPDFHILWEEPMYAIHAI 202	CC	-1- COPACTOR: Binds 1 2Fe-2S cluster and 3 4Fe-4S clusters per subunit
		CC	(By similarity).
		CC	-1- SIMILARITY: Belongs to the complex I 75 kDa subunit family.
RESULT 3		CC	-----
Q7Y3V5	PRELIMINARY; PRT; 357 AA.	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
AC	Q7Y3V5;	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
DT	01-OCT-2003 (TrEMBLrel. 25, Created)	CC	the European Bioinformatics Institute. There are no restrictions on its
DT	01-OCT-2003 (TrEMBLrel. 25, Last sequence update)	CC	use by non-profit institutions as long as its content is in no way
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)	CC	modified and this statement is not removed. Usage by and for commercial
DE	DNA adenine-methylase.	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
OS	Bacteriophage PY54.	CC	or send an email to license@isb-sib.ch).
OC	Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.	CC	-----
OX	NCBI_TaxID=172667;	DR	EMBL; BX248345; CAD95267.1; -
RN	[1]	DR	InterPro; IPR000283; Complex1_75K.
RP	SEQUENCE FROM N.A.	DR	InterPro; IPR001041; Ferredoxin.
RX	PubMed=12899832;	DR	InterPro; IPR006656; Molybdopterin.
RA	Hertwig S., Klein I., Schmidt V., Beck S., Hammerl J.A., Appel B.;	DR	InterPro; IPR006963; Molybdop_Fe4S4.
RT	"Sequence analysis of the genome of the temperate Versinia	DR	InterPro; IPR006657; Mol_dinuc_bind.
RL	enterocolitica phage PY54.";	DR	InterPro; IPR010228; NuOG.
RL	J. Mol. Biol. 331:605-622(2003).	DR	Pfam; PF00111; Fer2; 1.
DR	EMBL; AJ564013; CAD91820.1; -	DR	Pfam; PF04879; Molybdop_Fe4S4; 1.
DR	GO; GO:0003677; F:DNA binding; IEA.	DR	Pfam; PF01588; Molybdop_Binding; 1.
DR	GO; GO:0008170; P:N-methyltransferase activity; IEA.	DR	TIGRFAMS; TIGR01973; NuOG; 1.
DR	GO; GO:0006306; P:DNA methylation; IEA.	DR	PROSITE; PS00641; COMPLEX1_75K_1; 1.
DR	InterPro; IPR002941; N6/N4_Mtase.	DR	PROSITE; PS00642; COMPLEX1_75K_2; 1.
DR	InterPro; IPR002052; N6_Mtase.	DR	PROSITE; PS00643; COMPLEX1_75K_3; 1.
DR	Pfam; PF01555; N6_N4_Mtase; 1.	KW	2Fe-2S; 4Fe-4S; Complete proteome; Iron; Iron-sulfur; Metal-binding;
DR	PROSITE; PS00092; N6_MTASE; UNKNOWN_1.	KW	NAD; Oxidoreductase; Quinone.
XQ	Methyltransferase.	FT	METAL 49
SQ	SEQUENCE 357 AA; 40867 MW; 2152FA5655D63927 CRC64;	FT	METAL 60
	Query Match 47.8%; Score 54; DB 2; Length 357;	FT	METAL 63
	Best Local Similarity 47.4%; Pred. No. 9.6;	FT	METAL 77
	Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;	FT	METAL 111
		FT	METAL 115
		FT	METAL 118
		FT	METAL 124
		FT	METAL 164
		FT	METAL 167
		FT	METAL 170
		FT	METAL 214
		FT	METAL 240
		FT	METAL 243
		FT	METAL 247
		FT	METAL 275
		SQ	SEQUENCE 806 AA; 85412 MW; D674D74FA65FF3E81 CRC64;
RESULT 4			-----
NUOG_MYCBO	STANDARD; PRT; 806 AA.		Query Match 46.9%; Score 53; DB 1; Length 806;
AC	P59962;		Best Local Similarity 50.0%; Pred. No. 32;
DT	29-MAR-2004 (Rel. 43, Created)		Matches 12; Conservative 3; Mismatches 7; Indels 2; Gaps 1;
DT	29-MAR-2004 (Rel. 43, Last sequence update)		
DT	05-JUL-2004 (Rel. 44, Last annotation update)		
DE	NADH-quinone oxidoreductase chain G (EC 1.6.99.5) (NADH dehydrogenase		
DE	I, chain G) (NDH-1, chain G).		
GN	Name=nuoG; OrderedLocusNames=MB3175;		

QY 1 VVDVDFIVW--LEEAVSDLHRL 22
 DB 769 VTDMPDRVWPLNSAGSVHRQL 792

RESULT 5
 NUOG MYCTU
 ID NUOG MYCTU STANDARD; PRT; 806 AA.
 AC P95175;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE NADH-quinone oxidoreductase chain G (EC 1.6.99.5) (NADH dehydrogenase I, chain G) (NDH-1, chain G).
 DE I, chain G) (NDH-1, chain G).
 GN Name=nuog; OrderedLocNames=Rv3151, MT3239; ORFNames=MT303A2.07c;
 OS Mycobacterium tuberculosis
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 EX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.W.,
 RA Harris D.E., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III,
 RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
 RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
 RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
 RA Rogers J., Rutter S., Seeger K., Skelton S., Squares R.,
 RA Sultson J.E., Taylor K., Whitthead S., Barrrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 EX MEDLINE=22208494; PubMed=12218036;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
 RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
 RA Salzberg S.L., Delcher A.L., Utterback T.R., Weidman J.F., Khouri H.M.,
 RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
 RA Fraser C.M.;
 RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL J. Bacteriol. 184:5479-5490 (2002).
 CC -!- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
 CC sulfur (Fe-S) centers, to quinones in the respiratory chain. The
 CC immediate electron acceptor for the enzyme in this species is
 CC believed to be menaquinone. Couples the redox reaction to proton
 CC translocation (for every two electrons transferred, four hydrogen
 CC ions are translocated across the cytoplasmic membrane), and thus
 CC conserves the redox energy in a proton gradient (by similarity).
 CC -!- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
 CC -!- COFACTOR: Binds 1 2Fe-2S cluster and 3 4Fe-4S clusters per subunit
 CC (By similarity).
 CC -!- SIMILARITY: Belongs to the complex I 75 kDa subunit family.

DR EMBL; BX842582; CAB06288.1; -
 DR EMBL; AE007138; AAK47578.1; ALT_INIT.
 DR PIR; H70647; H70647.
 DR TIGR; MT3239; -
 DR TubercuList; Rv3151; -
 DR InterPro; IPR001450; 4Fe4S_ferredoxin.

DR InterPro; IPR000283; Complex1_75K.
 DR InterPro; IPR001041; Ferredoxin.
 DR InterPro; IPR006556; Molybdopterin.
 DR InterPro; IPR006963; Molybdop_Fe4S4.
 DR InterPro; IPR006657; Mol_dinuc_bind.
 DR InterPro; IPR010228; NuOG.
 DR Pfam; PF00111; Fer2; 1.
 DR Pfam; PF00384; Molybdopterin; 1.
 DR Pfam; PF04879; Molybdop_Fe4S4; 1.
 DR Pfam; PF01568; Molybdop_Binding; 1.
 DR TIGR; TIGR00353; 4Fe4S_Ferredoxin.
 DR TIGR; TIGR01973; NuOG; 1.
 DR PROSITE; PS00641; COMPLEX1_75K_1; 1.
 DR PROSITE; PS00642; COMPLEX1_75K_2; 1.
 DR PROSITE; PS00643; COMPLEX1_75K_3; 1.
 KW 2Fe-2S; 4Fe-4S; Complete proteome; Iron; Iron-sulfur; Metal-binding;
 KW NAD; Oxidoreductase; Quinone.
 FT METAL 49 49 Iron-sulfur 1 (2Fe-2S) (By similarity).
 FT METAL 60 60 Iron-sulfur 1 (2Fe-2S) (By similarity).
 FT METAL 63 63 Iron-sulfur 1 (2Fe-2S) (By similarity).
 FT METAL 77 77 Iron-sulfur 1 (2Fe-2S) (By similarity).
 FT METAL 111 111 Iron-sulfur 2 (4Fe-4S) (By similarity).
 FT METAL 115 115 Iron-sulfur 2 (4Fe-4S) (By similarity).
 FT METAL 118 118 Iron-sulfur 2 (4Fe-4S) (By similarity).
 FT METAL 124 124 Iron-sulfur 2 (4Fe-4S) (By similarity).
 FT METAL 164 164 Iron-sulfur 3 (4Fe-4S) (By similarity).
 FT METAL 167 167 Iron-sulfur 3 (4Fe-4S) (By similarity).
 FT METAL 170 170 Iron-sulfur 3 (4Fe-4S) (By similarity).
 FT METAL 214 214 Iron-sulfur 3 (4Fe-4S) (By similarity).
 FT METAL 240 240 Iron-sulfur 4 (4Fe-4S) (Potential).
 FT METAL 243 243 Iron-sulfur 4 (4Fe-4S) (Potential).
 FT METAL 247 247 Iron-sulfur 4 (4Fe-4S) (Potential).
 FT METAL 275 275 Iron-sulfur 4 (4Fe-4S) (Potential).
 FT METAL 474 474 I -> M (in Ref. 2).
 FT CONFLICT 604 604 T -> A (in Ref. 2).
 SQ SEQUENCE 806 AA; 85423 MW; 519A1EA83181064 CRC64;

Query Match 46.9%; Score 53; DB 1; Length 806;
 Best Local Similarity 50.0%; Pred. NO. 32;
 Matches 12; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

QY 1 VVDVDFIVW--LEEAVSDLHRL 22
 DB 769 VTDMPDRVWPLNSAGSVHRQL 792

RESULT 6
 Q6UDW3 PRELIMINARY; PRT; 3094 AA.
 AC Q6UDW3;
 DT 05-JUL-2004 (T-EMBLrel. 27, Created)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
 DE Erythrocyte membrane protein 1.
 GN Name=var;
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Malayan-Camp;
 RX PubMed=14651636;
 RA Kraemer S.M., Smith J.D.;
 RT "Evidence for the importance of genetic structuring to the structural
 RT and functional specialization of the Plasmodium falciparum var gene
 RT family."
 RT Mol. Microbiol. 50:1527-1538 (2003).
 DR EMBL; AV372127; AAQ73930.1; -
 DR InterPro; IPR000875; Cecropin.
 DR PROSITE; PS00268; CECROPIN; UNKNOWN 1.
 SQ SEQUENCE 3094 AA; 359444 MW; 31C0DA3CB5C17A07 CRC64;

Query Match 46.0%; Score 52; DB 2; Length 3094;

```

Best Local Similarity 52.9%; Pred. No. 1.9e+02;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 VPDFIVWLEEAASDLHR 20
DB 273 VQFLRWLTWIEDLYR 289

RESULT 7
AAQ73930 PRELIMINARY; PRT; 3094 AA.
AC AAQ73930;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Erythrocyte membrane protein 1.
GN VAR.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Malayan-Camp;
RX PubMed=14651636;
RA Kraemer S.M., Smith J.D.;
RT Evidence for the importance of genetic structuring to the structural
RT and functional specialization of the Plasmodium falciparum var gene
RT family."
RL Mol. Microbiol. 50:1527-1538(2003).
RD EMBL; AY372127; AAQ73930.1; -.
SQ SEQUENCE 3094 AA; 359444 MW; 31C0DA3CB5C17A07 CRC64;

Query Match 46.0%; Score 52; DB 2; Length 3094;
Best Local Similarity 52.9%; Pred. No. 1.9e+02;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 VPDFIVWLEEAASDLHR 20
DB 273 VQFLRWLTWIEDLYR 289

RESULT 8
Q9XX66 PRELIMINARY; PRT; 349 AA.
ID Q9XX66;
AC Q9XX66;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein Y102A5C.31.
GN Name=Y102A5C.31;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
OC Rhabditidae; Beloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gardner A.E.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RD EMBL; AL031627; CAA20969.1; -.
DR FIR; T26367; T26367.
DR WormPep; Y102A5C.31; CE20402.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004930; F:G-protein coupled receptor activity; IEA.
DR InterPro; IPR003003; 7TM_chemrecept2.

```

```

DR InterPro; IPR000168; Nm7TM_chemrecept.
DR Pfam; PF01604; 7tm5; 1.
KW Hypothetical protein.
SQ SEQUENCE 349 AA; 40149 MW; 64664F68B6D98DD5 CRC64;

Query Match 45.1%; Score 51; DB 2; Length 349;
Best Local Similarity 61.5%; Pred. No. 27;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVDVPDFIVWLEE 13
DB 180 IIDDPDFHLWLEE 192

RESULT 9
Q7PZJ5 PRELIMINARY; PRT; 894 AA.
ID Q7PZJ5;
AC Q7PZJ5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP9332 (Fragment).
GN Name=agCG54448; ORFNames=ENSANGG00000012809;
OS Anopheles gambiae str. PBST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB0100896; EAA00658.1; -.
DR InterPro; IPR003409; MORN.
DR Pfam; PF02493; MORN; 5.
FT NON TER 1
FT NON TER 1
SQ SEQUENCE 894 AA; 101315 MW; D2A93539D3653B79 CRC64;

Query Match 45.1%; Score 51; DB 2; Length 894;
Best Local Similarity 47.6%; Pred. No. 73;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 VVDVPDFIVWLEEAASDLHRA 21
DB 112 VVQYPLNLVWITVSELSHRS 132

RESULT 10
Q6RFG9 PRELIMINARY; PRT; 227 AA.
ID Q6RFG9;
AC Q6RFG9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit II (fragment).
GN Name=Cox2;
OS Pycnococcus provasolii.
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae; Mamiellales;
OC Pycnococcaceae; Pycnococcus.
OX NCBI_TaxID=41880;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14761653;
RA Hackett J.D., Yoon H.S., Soares M.B., Bonaldo M.F., Casavant T.L.,
RA Schertz T.E., Nosenko T., Bhattacharya D.;
RT "Migration of the plastid genome to the nucleus in a peridinin
RT dinoflagellate."
RL Curr. Biol. 14:213-218(2004).
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory

```

chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. Subunit 2 transfers the electrons from cytochrome c via its binuclear copper A center to the bimetallic center of the catalytic subunit 1 (By similarity).

-1- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome c + 2 H(2)O.

-1- COPACTOR: Copper A (By similarity).

-1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).

-1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family. EMBL: AY500835; AAR99499.1; -. GO: GO:0005739; C:mitochondrion; IEA. InterPro: IPR001505; Copper CuA. InterPro: IPR008972; Cupredoxin. InterPro: IPR002429; Cyt_c_ox_2. Pfam: PF00116; COX2; 1. Pfam: PF02790; COX2_TM; 1. PRINTS: PR01166; CYCOXIDASEII. ProDom: PD000131; Copper CuA; 1. Copper; Electron transport; Inner membrane; Membrane; Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane; Transport. NON_TER 1

SEQUENCE 227 AA; 25590 MW; CF7637C75DB2899B CRC64; 1

Query Match 44.2%; Score 50; DB 2; Length 227;
Best Local Similarity 47.1%; Pred. No. 25;
Matches 8; Conservative 4; Mismatches 5; Indels

Qy	2	VDPDFIVWLEEA	VSDL	18
			:	:
Db	207	VSPDDEIIWLEAKI	SEI	22

RESULT 11

AA999499	PRELIMINARY;	PRT;	227 AA.
ID	AA999499		
AC	AA999499;		
DT	29-MAR-2004 (TREMELrel. 27, Created)		
DT	29-MAR-2004 (TREMELrel. 27, Last sequence update)		
DT	29-MAR-2004 (TREMELrel. 27, Last annotation update)		
DE	DE Cytochrome oxidase subunit II (Fragment).		
DE	COX2.		
GN	Pycnococcus provasolii.		
OS	Pycnococcus provasolii.		
OG	Mitochondrion.		
OC	Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae; Mamiellales;		
OC	Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae; Mamiellales;		
OC	Pycnococcaceae; Pycnococcus.		
OX	NCBI_TaxID=41890;		
OX	[1]		
RN	SEQUENCE FROM N.A.		
RP			
RX	PubMed=14761653;		
RA	Hackett J.D., Yoon H.S., Soares M.B., Bernaldo M.F., Casavant T.L.,		
RA	Scheetz T.E., Nosenko T., Bhattacharya D.;		
RT	"Migration of the plastid genome to the nucleus in a peridinin		
RT	dinoflagellate.";		
RL	Curr. Biol. 14:213-218 (2004).		
DR	EMBL; AY500835; AA999499.1; -.		
KW	Mitochondrion.		
FT	NON TER	1	
FT	NON TER	1	
SQ	SEQUENCE	227 AA;	25590 MW; CF7637C75DB2899B CRC64;

Query Match 44.2%; Score 50; DB 2; Length 227;
Best Local Similarity 47.1%; Pred. No. 25;
Matches 8; Conservative 4; Mismatches 5; Indels

Qy	2	VDVPDFIVWLEEA	VSDL	18
			:	:
Db	207	VSPDDFIIWLEAK	ISEI	22

RESULT 12

Q8XF74 PRELIMINARY; PRT; 294 AA.

AC	QBXF74; Q7AM41;
DT	01-MAR-2002 (TrEMBLrel. 20, Created)
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT	01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE	Putative adenine-specific DNA-modification methylase.
GN	Name=yhdJ; OrderedLocusNames=STV3566, t3301;
OS	Salmonella typhi.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC	Enterobacteriaceae; Salmonella.
CC	NCBI_Taxid=601;
FN	[1]
FN	SEQUENCE FROM N.A.
RP	STRAIN=Ty2 / ATCC 700931;
RC	MEDLINE=25231367; PubMed=12644504;
RC	Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA	Burland V., Kodymiani V., Schwartz D.C., Blattner F.R.,
RA	"Comparative Genomics of Salmonella enterica serovar Typhi strains Ty2
RT	and Crl8.";
RT	J. Bacteriol. 185:2330-2337(2003).
RL	[2]
FN	SEQUENCE FROM N.A.
RC	STRAIN=Crl8;
RC	MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
EX	Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA	Churche C.M., Mungall K.B., Bentley S.D., Holden M.T.G., Sebahia M.,
RA	Baker S., Basham D., Brooks K.B., Chillingworth T., Connerton P.,
RA	Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA	Feltwell T., Hanlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA	Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA	Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA	Whitehead S., Barrall B.G.;
RT	"Complete genome sequence of a multiple drug resistant Salmonella
RT	enterica serovar Typhi Crl8.";
RL	Nature 413:848-852(2001).
RL	EMBL; AE016845; AA070836.1; -
DR	EMBL; AL627278; CAD07901.1; -
DR	HSSP; P23192; IG60
DR	GO; GO:0003677; F:DNA binding; IEA.
DR	GO; GO:0004868; F:methyltransferase activity; IEA.
DR	GO; GO:0008170; F:methyltransferase activity; IEA.
DR	GO; GO:0006306; P:DNA methylation; IEA.
DR	InterPro; IPR002295; D21N6 mtfase.
DR	InterPro; IPR001091; Met trans-GN4.
DR	InterPro; IPR002941; N6/N4 MtaSe.
DR	InterPro; IPR002052; N6 MtaSe.
DR	Pfam; PF01555; N6_N4 MtaSe; 1.
DR	PRINTS; PR00506; D21N6MTFRASE.
DR	PRINTS; PR00508; S21N4MTFRASE.
DR	PROSITE; PS00092; N6_MTASe; UNKNOWN 1.
DR	Complete proteome; Methyltransferase.
KO	SEQUENCE 294 AA; 33554 MW; C674376A21C04285 CRC64;

Query Match 44.2%; Score 50; DB 2; Length 294;
Best Local Similarity 40.0%; Pred. No. 33;
Matches 8: Conservative 3; Mismatches 9: Indels

QY	3	DVPDFIVWLEEAVSDLHRAL	22
		: : : : :	
DB	57	DEASEFLAWLYECIDECHRVL	76

RESULT 13

Q7CPL9 ID Q7CPL9 PRELIMINARY; PRT; 294 AA.

AC	Q/CPUS?	05-JUL-2004	(TrEMBLrel. 27, Created)
DT		05-JUL-2004	(TrEMBLrel. 27, Last sequence update)
DT		05-JUL-2004	(TrEMBLrel. 27, Last annotation update)
DT		05-JUL-2004	(TrEMBLrel. 27, Last annotation update)
DE			Putative methyltransferase (EC 2.1.1.72)

GN. *name=ymuo*; *Order=eloc*; *names=slm3366*;
OS. *Salmonella typhimurium*.
OC. Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC. Enterobacteriaceae; Salmonella.

```

OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=211534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
DR EMBL; AE008856; AAL22255.1; -
DR GO; GO:0008168; F:methyltransferase activity; IEA.
DR GO; GO:0009007; P:site-specific DNA-methyltransferase (adenin. .; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR InterPro; IPR002295; D2IN6_mtfrase.
DR InterPro; IPR001091; MetTfns_CN4.
DR InterPro; IPR002941; N6/N4_Mtase.
DR InterPro; IPR002052; N6_Mtase.
DR Pfam; PF01555; N6_N4_Mtase; 1.
DR PRINTS; PRO0508; D2IN6MTFRASE.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN1.
KW Complete proteome; Methyltransferase; Transferase.
SQ SEQUENCE 294 AA; 33554 MW; C674376A21C04285 CRC64;

Query Match 44.2%; Score 50; DB 2; Length 294;
Best Local Similarity 40.0%; Pred. No. 33;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 3 DVPDFIVMLEEAVSLHRL 22
DB 57 DEASFLAWLYECIDCHRYL 76

RESULT 14
AROB AGRT5
ID AROB AGRT5 STANDARD; PRT; 377 AA.
AC Q8U9V0;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE 3-dehydroquininate synthase (EC 4.2.3.4).
GN Name=aroB; OrderedLocustNames=Atu3625; AGR L 2401;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OC NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Bisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chunley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA Goodner B., Hinkle G., Gatungu S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houniel K., Gordon J., Vaudin M., Iatichouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,

```

```

RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
CC -1- CATALYTIC ACTIVITY: 3-deoxy-arabino-heptulosonate 7-phosphate = 3-
CC dehydroquininate + phosphate.
CC -1- COFACTOR: NAD and a divalent metal cation (By similarity).
CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC second step.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: Belongs to the dehydroquininate synthase family.
CC
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE009291; AAL44437.1; ALT INIT.
CC EMBL; AE008321; AAK89771.1; -.
CC PIR; A9281; A99281.
CC HSSP; P07547; 1NVE.
CC HAMAP; MF_00110; -.
CC InterPro; IPR002658; DHQ_synthase.
CC Pfam; PF01761; DHQ_synthase; 1.
CC TIGRFAMs; TIGR01357; aroB; 1.
KW Aromatic amino acid biosynthesis; Complete proteome; Lyase; NAD.
SQ SEQUENCE 377 AA; 40050 MW; 5B75B42CE204EA17 CRC64;

Query Match 44.2%; Score 50; DB 1; Length 377;
Best Local Similarity 44.4%; Pred. No. 42;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 VVDVPDFIVMLEEAVSDL 18
DB 200 LIDKPDFFWLEKWNDDI 217

RESULT 15
Q728W5 PRELIMINARY; PRT; 395 AA.
ID Q728W5
AC Q728W5;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=DVU2487;
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OS 8303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OC NCBI_TaxID=882;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15077118; DOI=10.1038/nbt959;
RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonel J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Davidson T.M., Zafar N., Zhou L., Radune D.,
RA Dmitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT Desulfovibrio vulgaris Hildenborough.";
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL; AS017317; AAS9659.1; -.
DR TIGR; DVU2487; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 395 AA; 43664 MW; 3B92825236EABEA CRC64;

```

Query Match 44.2%; Score 50; DB 2; Length 395;
 Best Local Similarity 50.0%; Pred. No. 45;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 7 FIVMLEEAVSDLHRAI 22
 |::||| |::|||
 Db 37 FVLMLESAPEDVRRRAI 52

Search completed: November 10, 2004, 14:50:32
 Job time : 31.3057 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 14.2755 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-62

Perfect score: 216

Sequence: 1 QRRNEQLDLADWLSSRRQGWMPNCRSRRGPG 39

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PTUS COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	61	28.2	218	4	US-09-252-991A-23272
2	61	28.2	369	4	US-09-489-039A-7324
3	61	28.2	1197	4	US-09-252-991A-30833
4	59	27.3	140	4	US-09-252-991A-19170
5	58	26.9	655	4	US-09-252-991A-22270
6	57.5	26.6	67	4	US-09-489-039A-9405
7	57	26.4	382	4	US-09-248-796A-18423
8	56	25.9	1226	1	US-08-280-443-2
9	56	25.9	1226	1	US-08-457-459-2
10	56	25.9	1226	1	US-08-555-678-2
11	56	25.9	1226	5	PCT-US95-02275-2
12	55	25.5	103	4	US-09-489-039A-7262
13	54.5	25.2	450	4	US-08-635-886C-188
14	54.5	25.2	450	4	US-08-974-690C-188
15	54.5	25.2	452	4	US-09-252-991A-20884
16	54	25.0	137	4	US-09-248-796A-28007
17	53	24.5	118	4	US-09-489-039A-11052
18	53	24.5	208	4	US-09-252-991A-23638
19	53	24.5	236	4	US-09-489-039A-9771
20	52	24.1	168	4	US-09-252-991A-19362
21	52	24.1	225	4	US-09-673-395A-480
22	52	24.1	402	4	US-09-252-991A-31178
23	51.5	23.8	483	4	US-09-252-991A-19224
24	51.5	23.8	1008	4	US-09-252-991A-19329
25	51	23.6	169	4	US-09-489-039A-9755
26	51	23.6	309	4	US-09-270-767-33777
27	51	23.6	309	4	US-09-270-767-48994

ALIGNMENTS

RESULT 1

US-09-252-991A-23272
Sequence 23272, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 23272

LENGTH: 218

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-23272

Query Match 28.2%; Score 61; DB 4; Length 218;

Best Local Similarity 48.3%; Pred. No. 0.64;

Matches 14; Conservative 0; Mismatches 5; Indels 10; Gaps 1;

QY 21 RSRQRGWMP-----NCRSRRGPG 39

Db 56 RDRGQWQYPCGATPARRRRRGPG 84

RESULT 2

US-09-489-039A-7324

Sequence 7324, Application US/09489039A

Patent No. 6610835

GENERAL INFORMATION:

APPLICANT: Gary Breton et. al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709.2004001

CURRENT APPLICATION NUMBER: US/09/489,039A

CURRENT FILING DATE: 2000-01-27

PRIOR APPLICATION NUMBER: US 60/117,747

PRIOR FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 14342

SEQ ID NO 7324

LENGTH: 369

TYPE: PRT

ORGANISM: Klebsiella pneumoniae

US-09-489-039A-7324

Sequence 4856, Ap
Sequence 28601, A
Sequence 17988, A
Sequence 32108, A
Sequence 27484, A
Sequence 23, Appl
Sequence 14177, A
Sequence 23992, A
Sequence 21510, A
Sequence 21865, A
Sequence 23882, A
Sequence 30524, A
Sequence 20214, A
Sequence 8590, Ap
Sequence 31776, A
Sequence 26751, A
Sequence 27110, A
Sequence 25751, A

479 4 US-09-543-681A-4856
586 4 US-09-252-991A-28601
629 4 US-09-252-991A-17988
151 4 US-09-252-991A-32108
290 4 US-09-252-991A-27484
305 4 US-10-162-012-23
398 4 US-09-489-039A-14177
444 4 US-09-252-991A-23992
582 4 US-09-252-991A-21510
849 4 US-09-252-991A-21865
938 4 US-09-252-991A-23882
171 4 US-09-252-991A-30524
275 4 US-09-252-991A-20214
308 4 US-09-489-039A-8590
417 4 US-09-252-991A-31776
431 4 US-09-252-991A-26751
446 4 US-09-252-991A-27110
503 4 US-09-252-991A-25751

28 51 23.6
29 51 23.6
30 51 23.6
31 50.5 23.4
32 50.5 23.4
33 50.5 23.4
34 50.5 23.4
35 50.5 23.4
36 50.5 23.4
37 50.5 23.4
38 50.5 23.4
39 50 23.1
40 50 23.1
41 50 23.1
42 50 23.1
43 50 23.1
44 50 23.1
45 50 23.1

Query Match 28.2%; Score 61; DB 4; Length 369;
Best Local Similarity 75.0%; Pred. No. 1.2; 1; Indels 0; Gaps 0;
Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 0;

QY 26 GQMPNCRSRG 37
DB 14 GWLPCRSRG 25
||:|||||

RESULT 3
US-09-252-991A-30833
; Sequence 30833, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30833
; LENGTH: 1197
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1028)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-30833

Query Match 28.2%; Score 61; DB 4; Length 1197;
Best Local Similarity 42.5%; Pred. No. 4.6; 7; Mismatches 12; Indels 4; Gaps 2;
Matches 17; Conservative 7; Mismatches 12; Indels 4; Gaps 2;

QY 2 RRGNEFQDLADAWLSSRRQRQWQMPNCR--SRRGPG 39
DB 908 RQDFEFLRDVAE--ALGRAGAGPRQPRRRVGRGPG 945
||:|||||:|:|:|

RESULT 4
US-09-252-991A-19170
; Sequence 19170, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19170
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19170

Query Match 27.3%; Score 59; DB 4; Length 140;
Best Local Similarity 53.6%; Pred. No. 0.74; 4; Mismatches 5; Indels 4; Gaps 2;
Matches 15; Conservative 4; Mismatches 5; Indels 4; Gaps 2;

QY 6 EFQLRDLRPSDEADLASREQRQWQL 74
DB 47 EFGELRDLRPSDEADLASREQRQWQL 74
||:|||||

RESULT 5
US-09-252-991A-22270
; Sequence 22270, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22270
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22270

Query Match 26.9%; Score 58; DB 4; Length 656;
Best Local Similarity 57.9%; Pred. No. 6.2; 8; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 21 RSRQRGQWQMPNCRSRRGPG 39
DB 138 RGRHAGCPPPGCRSRSGPG 156
||:|||||

RESULT 6
US-09-489-039A-9405
; Sequence 9405, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9405
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9405

Query Match 26.6%; Score 57.5; DB 4; Length 67;
Best Local Similarity 42.3%; Pred. No. 0.5; 6; Indels 5; Gaps 1;
Matches 11; Conservative 4; Mismatches 6; Indels 5; Gaps 1;

QY 16 WDLSSRSRQRGW-----QMPNCRSRR 36
DB 3 WSAWSSASRSRQWTVYRAVMPSCRSRR 28
||:|||||

RESULT 7
US-09-248-796A-18423
; Sequence 18423, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
US-09-248-796A-18423


```

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/555,678
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,794
; FILING DATE: 17-FEB-1994
; PRIOR APPLICATION NUMBER: US 08/280,443
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,459
; FILING DATE: 01-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST49DUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1226 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-555-678-2

Query Match 25.9%; Score 56; DB 1; Length 1226;
Best Local Similarity 36.1%; Pred. No. 25;
Matches 13; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

QY 3 RGNFQLRLADAWDLSSRSRQGMQMPNCRSRG 38
DB 84 RGRQVDIRGVRGVHLGSQLQGRGFQHPSPRGSLP 119

RESULT 11
PCT-US95-02275-2
; Sequence 2, Application PC/TUS9502275
; GENERAL INFORMATION:
; APPLICANT: Wistar Institute of Anatomy & Biology
; TITLE OF INVENTION: RNA Editing Enzyme and Methods of Use
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Howson and Howson
; STREET: Spring House Corporate Cntr, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02275
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/280,443
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,794
; FILING DATE: 17-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215

```

```

;
; REFERENCE/DOCKET NUMBER: WST49BPCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1226 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-02275-2

Query Match 25.9%; Score 56; DB 5; Length 1226;
Best Local Similarity 36.1%; Pred. No. 25;
Matches 13; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

QY 3 RGNFQLRLADAWDLSSRSRQGMQMPNCRSRG 38
DB 84 RGRQVDIRGVRGVHLGSQLQGRGFQHPSPRGSLP 119

RESULT 12
US-09-489-039A-7262
; Sequence 7262, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7262
; TYPE: PRT
; LENGTH: 103
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-7262

Query Match 25.5%; Score 55; DB 4; Length 103;
Best Local Similarity 54.5%; Pred. No. 1.9;
Matches 12; Conservative 3; Mismatches 3; Indels 4; Gaps 2;

QY 16 WDLSSRSRQGMQMPNCRSRG 37
DB 22 WAISNRS---GWRMP-CASRSG 39

RESULT 13
US-08-635-886C-188
; Sequence 188, Application US/08635886C
; Patent No. 6555114
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2752-18
; CURRENT APPLICATION NUMBER: US/08/635,886C
; CURRENT FILING DATE: 1996-04-25
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 188
; LENGTH: 450
; TYPE: PRT
; ORGANISM: hepatitis C virus

```

US-08-635-886C-188

Query Match 25.2%; Score 54.5; DB 4; Length 450;
Best Local Similarity 36.8%; Pred. No. 13;
Matches 14; Conservative 5; Mismatches 16; Indels 1

Qy 2 RRGNEFQLRLDADAWDLSSRSRQRGWQMPCNCRSRRPGP 39
||| : ||| : ||| :
Dd 39 RRGEPLGVRATRK-T--SERSQPRGWROIPKARRPEG 73

RESULT 14

```

US-08-974-690C-188
; Sequence 188, Application US/08974690C
; Patent No. 661333
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/08/974,690C
; CURRENT FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 188
; LENGTH: 450
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-974-690C-188

```

Query Match 25.2%; Score 54.5; DB 4; Length 450;
Best Local Similarity 36.8%; Pred. No. 13;
Matches 14; Conservative 5; Mismatches 16; Indels

QY 2 RRGNEFQLRDLDADWDLSRSRQRQGWPNCRRSRPGP
||| : ||| | | | | : |
D6 39 RGPELGVATRKT--SERSOPRGWROI PKARRPEG 73

RESIT.T 15

```

US-09-252-991A-20884
; Sequence 20884, Application US/09252991A
; Patent No. 6551795
;
GENERAL INFORMATION:
;
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20884
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20884

```

```
Query Match      25.2%; Score 54.5; DB 4; Length 452;
Best Local Similarity 35.3%; Pred. No. 13;
Matches 18; Conservative 5; Mismatches 11; Indels 1
```

QY 2 RRGNE---FQLRDLADAWDLSSRSRQRGWQMPNCR-----SRRG 37

381 RRTSERPGTSLRRTRPA--LPRRTRRRWWORPACRGLPLBPSSSGRVARRG 429

Search completed: November 10, 2004, 14:55:44
Job time : 15.3255 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 45.1075 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-62

Perfect score: 216

Sequence: 1 QRRGNEFQRLDLADAWLSSRRQGWQMPNCRSRGPG 39

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 35322586 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09C_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	216	100.0	39	14 US-10-092-750-62	Sequence 62, Appl
2	64.5	29.9	286	15 US-10-425-114-66943	Sequence 66943, A
3	64.5	29.9	391	17 US-10-425-115-36744	Sequence 336744, A
4	61.5	28.5	138	17 US-10-425-115-242951	Sequence 242951, A
5	60.5	28.0	88	9 US-09-764-877-1924	Sequence 1924, Ap
6	60.5	28.0	88	15 US-10-242-515-1924	Sequence 1924, Ap
7	60.5	28.0	141	15 US-10-425-114-67835	Sequence 67835, A
8	60.5	28.0	498	17 US-10-425-115-209098	Sequence 209098, A
9	60	27.8	76	17 US-10-425-115-225012	Sequence 225012, A
10	60	27.8	190	17 US-10-425-115-36562	Sequence 36562, A
11	58	26.9	168	14 US-10-369-493-11976	Sequence 11976, A
12	56.5	26.2	78	16 US-10-437-963-142826	Sequence 142826, A
13	56.5	26.2	551	16 US-10-437-963-110875	Sequence 110875, A

14	56.5	26.2	551	16	US-10-437-963-123723	Sequence 123723, A
15	56	25.9	95	17	US-10-425-115-366113	Sequence 366113, A
16	56	25.9	716	15	US-10-412-699B-404	Sequence 404, App
17	56	25.9	1425	14	US-10-233-553-20	Sequence 20, Appl
18	55	25.5	136	15	US-10-282-122A-53966	Sequence 53966, A
19	55	25.5	877	14	US-10-369-493-19266	Sequence 19266, A
20	55	25.5	913	16	US-10-437-963-197298	Sequence 197298, A
21	54.5	25.2	450	15	US-10-651-185-188	Sequence 188, App
22	54	25.0	160	16	US-10-437-963-175669	Sequence 175669, A
23	54	25.0	210	17	US-10-425-115-186600	Sequence 186600, A
24	54	25.0	1623	16	US-10-437-963-202524	Sequence 202524, A
25	53.5	24.8	164	16	US-10-437-963-127029	Sequence 127029, A
26	53.5	24.8	169	14	US-10-156-761-7602	Sequence 7602, Ap
27	53	24.5	91	17	US-10-425-115-342590	Sequence 342590, A
28	53	24.5	160	17	US-10-425-115-218695	Sequence 218695, A
29	53	24.5	201	16	US-10-437-963-138893	Sequence 138893, A
30	53	24.5	224	17	US-10-425-115-264161	Sequence 264161, A
31	53	24.5	231	15	US-10-282-122A-59841	Sequence 59841, A
32	52.5	24.3	99	9	US-09-864-761-48852	Sequence 48852, A
33	52.5	24.3	138	16	US-10-767-701-53818	Sequence 53818, A
34	52.5	24.3	163	17	US-10-425-115-264163	Sequence 264163, A
35	52.5	24.3	210	15	US-10-425-114-70467	Sequence 70467, A
36	52.5	24.3	442	15	US-10-425-114-43237	Sequence 43237, A
37	52.5	24.3	581	15	US-10-425-114-62613	Sequence 62613, A
38	52	24.1	117	15	US-10-425-114-47996	Sequence 47996, A
39	52	24.1	137	16	US-10-437-963-189043	Sequence 189043, A
40	52	24.1	149	16	US-10-437-963-115114	Sequence 115114, A
41	52	24.1	183	16	US-10-437-963-108992	Sequence 108992, A
42	52	24.1	221	15	US-10-425-114-49030	Sequence 49030, A
43	52	24.1	368	15	US-10-424-599-151206	Sequence 151206, A
44	52	24.1	368	15	US-10-425-114-50824	Sequence 50824, A
45	52	24.1	533	14	US-10-156-761-13934	Sequence 13934, A

ALIGNMENTS

RESULT 1
US-10-092-750-62
; Sequence 62, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-62

Query Match 100.0%; Score 216; DB 14; Length 39;
Best Local Similarity 100.0%; Pred. No. 4e+21;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRRGNEFQRLDLADAWLSSRRQGWQMPNCRSRGPG 39
DB 1 QRRGNEFQRLDLADAWLSSRRQGWQMPNCRSRGPG 39
RESULT 2
US-10-425-114-66943
; Sequence 66943, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:


```

; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1924
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (72)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (77)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (84)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-242-515-1924

Query Match      28.0%; Score 60.5; DB 15; Length 88;
Best Local Similarity 45.9%; Pred. No.3;
Matches 17; Conservative 1; Mismatches 10; Indels 9; Gaps 3;

Qy      12 LADAWD-LSSRSRORG-----WQMPN---CRSRRGPG 39
Db      25 LLPRWPLSSRXRSSGCGCRWWSGFSRWQCRRSRSPG 61

RESULT 7
US-10-425-114-67835
; Sequence 67835, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128

```

; OTHER INFORMATION: Clone ID: MRT4577_136802C.1.pep
US-10-425-115-225012

Query Match 27.8%; Score 60; DB 17; Length 76;
Best Local Similarity 63.2%; Pred. No. 3;
Matches 12; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 18 LSSRSRQSGWQWPNCRSR 36
DB 24 LSVRSRRRWPAAPCRSTR 42

RESULT 10

US-10-425-115-361562
; Sequence 361562, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 361562
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(190)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_92922C.1.pep
US-10-425-115-361562

Query Match 27.8%; Score 60; DB 17; Length 190;
Best Local Similarity 50.0%; Pred. No. 7.6;
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 19 SRSRQSGWQWPNCRSR 38
DB 89 TTSTRSGWPNPACRSRTP 108

RESULT 11

US-10-369-493-11976
; Sequence 11976, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Cher, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11976
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Mesorhizobium loti
US-10-369-493-11976

Query Match 26.9%; Score 58; DB 14; Length 168;

Best Local Similarity 44.4%; Pred. No. 12;
Matches 12; Conservative 3; Mismatches 8; Indels 4; Gaps 1;

QY 7 FQRLDLADAWDLS-----SRSRQSGWQ 29
DB 118 FNKSLADRWISDGDIDNLRQLGWOI 144

RESULT 12

US-10-437-963-142826
; Sequence 142826, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 142826
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_43795C.1.pep
US-10-437-963-142826

Query Match 26.2%; Score 56.5; DB 16; Length 78;
Best Local Similarity 50.0%; Pred. No. 9;
Matches 13; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

QY 12 LADWDLSSRQSGWQWPNCRSR 37
DB 17 LASAWTSSRRRAQGC-LPECRWKG 41

RESULT 13

US-10-437-963-110875
; Sequence 110875, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 110875
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_14902C.1.pep
US-10-437-963-110875

Query Match 26.2%; Score 56.5; DB 16; Length 551;
Best Local Similarity 37.1%; Pred. No. 64;

Search completed: November 11, 2004, 02:43:18
Job time : 46.1575 secs

Matches 13; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

QY 2 REGNEFQRLDADAWDLSSRSRQGMPCNCRSR 36
DB 510 RYRSSELRDQKRWERATR-EDRAWVDFARPR 543

RESULT 14

US-10-437-963-123723
; Sequence 123723, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 123723
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_2652C.1.pap
US-10-437-963-123723

Query Match 26.2%; Score 56.5; DB 16; Length 551;
Best Local Similarity 39.3%; Pred. No. 64;
Matches 11; Conservative 4; Mismatches 10; Indels 3; Gaps 1;

QY 10 RDLADAWDLSSRSRQGMPCNCRSR 34
DB 20 REAAAWQFAGNHRRRGWGWLPRCAS 47

RESULT 15

US-10-425-115-366113
; Sequence 366113, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 366113
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_97060C.1.pap
US-10-425-115-366113

Query Match 25.9%; Score 56; DB 17; Length 95;
Best Local Similarity 35.3%; Pred. No. 13;
Matches 12; Conservative 4; Mismatches 18; Indels 0; Gaps 0;

QY 4 GNEFQRLDADAWDLSSRSRQGMPCNCRSRG 37
DB 47 GEQWQKYLGAICSCFCFGGWRGWRCDNCRPRG 80

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 9.56604 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-62

Perfect score: 216

Sequence: 1 QRGNEFQLRLADAMDLSSRSRQRGMPCNCRRRGPG 39

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64	29.6	466	1 B43332	glutamate decarboxylase
2	64	29.6	466	1 S24234	glutamate decarboxylase
3	64	29.6	466	2 C85726	glutamate decarboxylase
4	64	29.6	466	2 B90691	glutamate decarboxylase
5	60	27.8	466	2 E91178	glutamate decarboxylase
6	60	27.8	466	2 F86024	glutamate decarboxylase
7	59	27.3	134	2 E82953	conserved hypother
8	57	26.4	462	2 AH1130	glutamate decarboxylase
9	57	26.4	464	2 AC1370	glutamate decarboxylase
10	57	26.4	464	2 AB1740	glutamate decarboxylase
11	56	25.9	739	2 H85245	glutamate decarboxylase
12	56	25.9	739	2 T05163	Vpi like protein [
13	56	25.9	1226	1 S65593	adenosine deaminas
14	56	25.9	2034	2 T22147	hypothetical prote
15	55	25.5	304	2 AE3623	hypothetical prote
16	55	25.5	3305	2 T18358	apolipoprotein prec
17	54.5	25.2	154	2 JQ1137	hypothetical prote
18	54.5	25.2	513	2 FC1284	rRNA methyltransf
19	54.5	25.2	194	2 C83312	genome polyprotein
20	52	24.1	194	2 C83312	hypothetical prote
21	52	24.1	746	2 T35811	probable phosphodi
22	52	24.1	1516	2 T01055	hypothetical prote
23	51.5	23.8	438	2 AD0437	conserved hypother
24	51.5	23.8	2605	2 T18552	safiramycin Mxi syn
25	51.5	23.6	345	2 AG3186	hypothetical prote
26	51	23.6	365	2 AH2606	transcription regu
27	51	23.6	365	2 G97388	conserved hypother
28	50.5	23.4	332	2 F75473	conserved hypother
29	50.5	23.4	621	2 E82768	conserved hypother

30 50.5 23.4 885 2 C83441 two-component sens
31 50 23.1 299 2 T43442 hypothetical prote
32 50 23.1 375 2 AB0725 ribonuclease D [im
33 50 23.1 946 2 T31488 hypothetical prote
34 50 23.1 1377 2 E86034 rnsA protein in th
35 50 23.1 1377 2 C65159 probable glucan sy
36 50 23.1 1510 2 C84727 hypothetical prote
37 50 23.1 1556 2 D36793 hypothetical prote
38 49.5 22.9 111 2 C72714 small T antigen -
39 49.5 22.9 195 1 TVVPA hypothetical prote
40 49.5 22.9 247 2 T08721 hypothetical prote
41 49.5 22.9 271 1 TVVPMP middle T antigen -
42 49.5 22.9 421 1 TVVPMP middle T antigen -
43 49.5 22.9 440 1 TVVPMP aldehyde-ferredoxi
44 49.5 22.9 621 2 A71221 aldehyde ferredoxi
45 49.5 22.9 624 2 E75188

ALIGNMENTS

RESULT 1

B43332
glutamate decarboxylase (EC 4.1.1.15) beta - Escherichia coli (strain K-12)
N;Alternate names: L-glutamate l-carboxy-lyase
C;Species: Escherichia coli
C;Date: 10-Mar-1994 #sequence_revision 23-Mar-1995 #text_change 09-Jul-2004
C;Accession: B43332; S30261; H64902
R;Smith, D.K.; Kassam, T.; Singh, B.; Elliott, J.F.
J. Bacteriol. 174, 5820-5826, 1992
A;Title: Escherichia coli has two homologous glutamate decarboxylase genes that map to
A;Reference number: A43332; MUID:92394884; PMID:1522060
A;Accession: B43332
A;Molecule type: DNA
A;Residues: 1-466 <SMI>
A;Cross-references: UNIPROT:P28302; GB:M84025; NID:G146059; PIDN:AAA23834.1; PID:G146059
R;Yoshida, T.; Ueguchi, C.; Yamada, H.; Mizuno, T.
Mol. Gen. Genet. 237, 113-122, 1993
A;Title: Function of the Escherichia coli nucleoid protein, H-NS: molecular analysis of
A;Reference number: S30261; MUID:93204884; PMID:8455549
A;Accession: S30261
A;Molecule type: protein
A;Residues: 1-4, 'LOVL', 7-15 <YOS>
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: H64902
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-466 <BLAT>
A;Cross-references: GSI:AE000246; GB:U000096; NID:G1787764; PIDN:AAC74566.1; PID:G1787764
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: gadB
A;Map position: 78 min
C;Function:
A;Description: catalyzes the formation of gamma-aminobutyric acid, from L-glutamic acid
A;Note: in E. coli, two isoforms (alpha and beta) have been found, each encoded by a se
C;Superfamily: Escherichia coli glutamate decarboxylase
C;Keywords: carbon-carbon lyase; carboxy-lyase; homohexamer; phosphoprotein; pyridoxal
F;276/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 29.6%; Score 64; DB 1; Length 466;

Best Local Similarity 50.0%; Pred. No. 0.82; Mismatches 4; Gaps 1;

Matches 14; Conservative 4; Indels 6; Indels 4; Gaps 1;

7 FQLRLAD----AWDLSSRSRQRGMPCNCRRRGPG 30

380 FKLKGDGDPGYTLVLDLSERLRLRGWQVP 407

```

RESULT 2
S24234
glutamate decarboxylase (EC 4.1.1.15) alpha - Escherichia coli (strain K-12)
N:Alternate names: GAD alpha protein; L-glutamate 1-carboxy-lyase
C:Species: Escherichia coli
C:Date: 22-Jan-1993 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: S47737; H65149; S24234; S23421; A43332; FN0616
R:Plunkett, G.
submitted to the EMBL Data Library, March 1994
A:Reference number: S47666
A:Accession: S47737
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-466 <PLU>
A:Cross-references: UNIPROT:P80063; EMBL:U00039; NID:G466582; PTDN:AA818493.1; PID:G4666
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: H65149
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-466 <BLAT>
A:Cross-references: GB:AB000428; GB:U00096; NID:G1789931; PIDN:AAC76542.1; PID:G1789934;
A:Experimental source: strain K-12, substrain MG1655
R:Maras, B.; Sweeney, G.; Barra, D.; Bossa, F.; John, R.A.
Eur. J. Biochem. 204, 93-98, 1992
A:Title: The amino acid sequence of glutamate decarboxylase from Escherichia coli. Evolu
A:Reference number: S23421; MUID:9215241; PMID:1740158
A:Accession: S24234
A:Molecule type: protein
A:Residues: 1-63,'S',65-72,'R',74-152,'N',154-164,'S',166-354,'N',356-466 <MAR>
A:Accession: S23421
A:Molecule type: DNA
A:Residues: 148-164,'S',166-207,'N',209-294,'V',296-431 <VAR1>
A:Cross-references: GB:X63123; NID:G41601; PIDN:CAA44834.1; PID:G938166
R:Smith, D.K.; Kassam, T.; Singh, B.; Elliott, J.F.
J. Bacteriol. 174, 5820-5826, 1992
A:Title: Escherichia coli has two homologous glutamate decarboxylase genes that map to d
A:Reference number: A43332; MUID:92394884; PMID:1522060
A:Accession: A43332
A:Molecule type: DNA
A:Residues: 1-466 <SMI>
A:Cross-references: GB:M84024; NID:G146057; PIDN:AA238333.1; PID:G146058
A:Note: sequence extracted from NCBI backbone (NCBIN:112979, NCBI:P:112980)
R:Yoshida, T.; Yamashino, T.; Ueguchi, C.; Mizuno, T.
Biosci. Biotechnol. Biochem. 57, 1568-1569, 1993
A:Title: Expression of the Escherichia coli dimorphic glutamic acid decarboxylases is re
A:Reference number: FN0616; MUID:940333862; PMID:7764225
A:Accession: FN0616
A:Molecule type: protein
A:Residues: 1-21,'X' <YOS>
C:Genetics:
A:Gene: Gada
A:Map position: 78 min
C:Function:
A:Description: catalyzes the formation of gamma-aminobutyric acid, from L-glutamic acid
A:Note: in E. coli, two isoforms (alpha and beta) have been found, each encoded by a sep
C:Superfamily: Escherichia coli glutamate decarboxylase
C:Keywords: carbon-carbon lyase; carboxy-lyase; homohexamer; phosphoprotein; pyridoxal P
F:276/Binding site: pyridoxal phosphate (lys) (covalent) #status not shown
Query Match 29.6%; Score 64; DB 1; Length 466;
Best Local Similarity 50.0%; Pred. No. 0.82;
Matches 14; Conservative 4; Mismatches 6; Indels 4; Gaps 1;

QY 7 FQRLDLAD----AWDLSSRSRQGWQMP 30
Db 380 FKLKGDGDPGYTYLDLSERLRLRGWQVP 407

RESULT 3
S24234
glutamate decarboxylase (EC 4.1.1.15) alpha - Escherichia coli (strain K-12)
N:Alternate names: GAD alpha protein; L-glutamate 1-carboxy-lyase
C:Species: Escherichia coli
C:Date: 22-Jan-1993 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: S47737; H65149; S24234; S23421; A43332; FN0616
R:Plunkett, G.
submitted to the EMBL Data Library, March 1994
A:Reference number: S47666
A:Accession: S47737
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-466 <PLU>
A:Cross-references: UNIPROT:P80063; EMBL:U00039; NID:G466582; PTDN:AA818493.1; PID:G4666
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: H65149
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-466 <BLAT>
A:Cross-references: GB:AB000428; GB:U00096; NID:G1789931; PIDN:AAC76542.1; PID:G1789934;
A:Experimental source: strain K-12, substrain MG1655
R:Maras, B.; Sweeney, G.; Barra, D.; Bossa, F.; John, R.A.
Eur. J. Biochem. 204, 93-98, 1992
A:Title: The amino acid sequence of glutamate decarboxylase from Escherichia coli. Evolu
A:Reference number: S23421; MUID:9215241; PMID:1740158
A:Accession: S24234
A:Molecule type: protein
A:Residues: 1-63,'S',65-72,'R',74-152,'N',154-164,'S',166-354,'N',356-466 <MAR>
A:Accession: S23421
A:Molecule type: DNA
A:Residues: 148-164,'S',166-207,'N',209-294,'V',296-431 <VAR1>
A:Cross-references: GB:X63123; NID:G41601; PIDN:CAA44834.1; PID:G938166
R:Smith, D.K.; Kassam, T.; Singh, B.; Elliott, J.F.
J. Bacteriol. 174, 5820-5826, 1992
A:Title: Escherichia coli has two homologous glutamate decarboxylase genes that map to d
A:Reference number: A43332; MUID:92394884; PMID:1522060
A:Accession: A43332
A:Molecule type: DNA
A:Residues: 1-466 <SMI>
A:Cross-references: GB:M84024; NID:G146057; PIDN:AA238333.1; PID:G146058
A:Note: sequence extracted from NCBI backbone (NCBIN:112979, NCBI:P:112980)
R:Yoshida, T.; Yamashino, T.; Ueguchi, C.; Mizuno, T.
Biosci. Biotechnol. Biochem. 57, 1568-1569, 1993
A:Title: Expression of the Escherichia coli dimorphic glutamic acid decarboxylases is re
A:Reference number: FN0616; MUID:940333862; PMID:7764225
A:Accession: FN0616
A:Molecule type: protein
A:Residues: 1-21,'X' <YOS>
C:Genetics:
A:Gene: Gada
A:Map position: 78 min
C:Function:
A:Description: catalyzes the formation of gamma-aminobutyric acid, from L-glutamic acid
A:Note: in E. coli, two isoforms (alpha and beta) have been found, each encoded by a sep
C:Superfamily: Escherichia coli glutamate decarboxylase
C:Keywords: carbon-carbon lyase; carboxy-lyase; homohexamer; phosphoprotein; pyridoxal P
F:276/Binding site: pyridoxal phosphate (lys) (covalent) #status not shown
Query Match 29.6%; Score 64; DB 1; Length 466;
Best Local Similarity 50.0%; Pred. No. 0.82;
Matches 14; Conservative 4; Mismatches 6; Indels 4; Gaps 1;

QY 7 FQRLDLAD----AWDLSSRSRQGWQMP 30
Db 380 FKLKGDGDPGYTYLDLSERLRLRGWQVP 407

RESULT 4
B90891
glutamate decarboxylase isozyme [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: B90891
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: B90891
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-466 <HAY>
A:Cross-references: UNIPROT:P28302; GB:BA000007; PIDN:BA835521.1; PID:G13361564; GSPDB:
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs2098
C:Superfamily: Escherichia coli glutamate decarboxylase
Query Match 29.6%; Score 64; DB 2; Length 466;
Best Local Similarity 50.0%; Pred. No. 0.82;
Matches 14; Conservative 4; Mismatches 6; Indels 4; Gaps 1;

QY 7 FQRLDLAD----AWDLSSRSRQGWQMP 30
Db 380 FKLKGDGDPGYTYLDLSERLRLRGWQVP 407

RESULT 5
E91178
glutamate decarboxylase isozyme [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: E91178
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: E91178
A:Status: preliminary
A:Molecule type: DNA

```

A;Residues: 1-466 <HAY>
A;Cross-references: UNIPROT:P58228; GB:BAQ00007; PIDN:BAB37820.1; PID:G13363871; GSPDB:G
A;Experimental source: strain O157:H7, substrain RMD 050952
C;Genetics:
A;Gene: ECs4397
C;Superfamily: Escherichia coli glutamate decarboxylase

Query Match 27.8%; Score 60; DB 2; Length 466;
Best Local Similarity 46.4%; Pred. No. 2.8;
Matches 13; Conservative 5; Mismatches 6; Indels 4; Gaps 1;

QY 7 FQLRDLDL----AWDLSSRSRQGWMP 30
Db 380 FKLKEGDPGVTLXDLSERLRGWQVP 407

RESULT 6
R86024
glutamate decarboxylase isozyme [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: F86024
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoumis, K.; Apodaca,
Nature 409, 528-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: AB5480; MUID:21074935; PMID:11206551
A;Accession: F86024
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-466 <STO>
A;Cross-references: UNIPROT:P58228; GB:AB005174; NID:G12518229; PIDN:AAQ58658.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EBL933
C;Genetics:
A;Gene: gadA
C;Superfamily: Escherichia coli glutamate decarboxylase

Query Match 27.8%; Score 60; DB 2; Length 466;
Best Local Similarity 46.4%; Pred. No. 2.8;
Matches 13; Conservative 5; Mismatches 6; Indels 4; Gaps 1;

QY 7 FQLRDLDL----AWDLSSRSRQGWMP 30
Db 380 FKLKEGDPGVTLXDLSERLRGWQVP 407

RESULT 7
B82953
conserved hypothetical protein PA5536 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: B82953
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A;Reference number: B82950; MUID:20437337; PMID:10984043
A;Accession: B82953
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-134 <STO>
A;Cross-references: UNIPROT:Q9HT38; GB:AB004966; GB:AE004091; NID:G9951872; PIDN:AAQ0892
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA5536
C;Superfamily: Escherichia coli dosage-dependent dnaK suppressor protein dksA

Query Match 27.3%; Score 59; DB 2; Length 134;
Best Local Similarity 53.6%; Pred. No. 1.1;
Matches 15; Conservative 4; Mismatches 5; Indels 4; Gaps 2;

QY 6 EF-QLRDL--ADAWDLSSRSRQGWQM 29

Db 41 EFGELDLRPSDEADLASREQRWQL 68

RESULT 8
AH1130
glutamate decarboxylase homolog lmo0447 [imported] - Listeria monocytogenes (strain EGD
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AH1130
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, F
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitouran, A.; M
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlanc
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1130
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-462 <GLA>
A;Cross-references: UNIPROT:Q9F5P3; GB:NC_003210; PIDN:CAC98526.1; PID:G16409824; GSPDE
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo0447
C;Superfamily: Escherichia coli glutamate decarboxylase

Query Match 26.4%; Score 57; DB 2; Length 462;
Best Local Similarity 42.9%; Pred. No. 7.1;
Matches 12; Conservative 8; Mismatches 4; Indels 4; Gaps 2;

QY 7 FQLRDLDL-AW---DLSSRSRQGWMP 30
Db 374 YKLKDDANVAVTLTDLADRLQMRGWQVP 401

RESULT 9
AC1370
glutamate decarboxylase homolog lmo2363 [imported] - Listeria monocytogenes (strain EGD
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AC1370
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitouran, A.; M
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlanc
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AC1370
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-464 <GLA>
A;Cross-references: UNIPROT:Q9EYW9; GB:NC_003210; PIDN:CAD00441.1; PID:G16411851; GSPDE
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo2363
C;Superfamily: Escherichia coli glutamate decarboxylase

Query Match 26.4%; Score 57; DB 2; Length 464;
Best Local Similarity 39.3%; Pred. No. 7.2;
Matches 11; Conservative 8; Mismatches 5; Indels 4; Gaps 1;

QY 7 FQLRDLDLADA----WDLSSRSRQGWMP 30
Db 376 YKLKDDANVAVTLTDLADRLQMRGWQVP 403

RESULT 10
AB1740
glutamate decarboxylase homolog lin2463 [imported] - Listeria innocua (strain Clp11262)


```

A;Gene: CESP:F43G9.6
A;Map position: 1
A;Introns: 23/3; 65/3; 163/3; 258/3; 330/3; 400/2; 437/3; 578/2; 706/3; 919/3; 952/2; 11/1
Query Match      25.9%   Score 56; DB 2; Length 2034;
Best Local Similarity 44.0%; Pred. No. 42;
Matches 11; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
Qy 13 ADAWDLSSRSRQRGWMPNCRSRRG 37
       :|::|||::|||
Db 1828 SDGVYSTKKRQGRGVKCPKCCTRRG 1852

RESULT 15
AE3623
Glucamate decarboxylase (EC 4.1.1.15) [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AE3623
R;DelVecchio, V.G.; Kapatal, R.J.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
R.; Mazur, M.; Veltman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessé
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AE3623
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-304 <KUR>
A;Cross-references: UNIPROT:Q8VBJO; GB:A5008918; PIDN:AAL54152.1; PID:g17985117; GSPDB:G
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI10910
A;Map position: II
C;Keywords: carbon-carbon lyase; carboxy-lyase

Query Match      25.5%   Score 55; DB 2; Length 304;
Best Local Similarity 42.9%; Pred. No. 8;
Matches 12; Conservative 5; Mismatches 7; Indels 4; Gaps 1;
Qy 7 FOURDIAD----AWDLSRSSRQRGWMP 30
       ||::|::|::|::|::|::|::|
Db 217 FRREGDPGYSLYDISERLRLTGWOVP 244

```

Search completed: November 10, 2004, 14:52:28
Job time : 9.56604 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 51.9509 Seconds
(without alignments)
431.938 Million cell updates/s

Title: US-10-092-750-62

Perfect score: 216
Sequence: 1 QRRGNEFQLRDLADAWDLSSRSRQRGWQMPNCRSRGPG 39

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Minimum DB seq length:	0
Maximum DB seq length:	2000000

Post-processing: Minimum Match of

Maximum Match 100%
Listing first 45 summaries

Database : UniProt 02:*

```

uniprot:
uniprot_sprot:
1: uniprot sprot:

```

```
2: uniprot_trembl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	64	29.6	466	1	DCEA_ECOLI	P80063	escherichia
2	64	29.6	466	1	DCBB_ECOLI	P28302	escherichia
3	64	29.6	466	2	Q7UAY1	Q7UAY1	shigella fl
4	64	29.6	487	2	Q83PR1	Q83PR1	shigella fl
5	64	29.6	489	2	Q8PHG5	Q8PHG5	escherichia
6	60.5	28.0	234	2	Q00833	Q00833	trypanosoma
7	60.5	28.0	269	2	Q8CBH3	Q8CBH3	mus musculus
8	60	27.8	220	2	Q6C9I7	Q6C9I7	yarrowia li
9	60	27.8	466	1	DCEA_ECO57	P58228	escherichia
10	59.5	27.5	730	2	Q81VF8	Q81VF8	homo sapiens
11	59	27.3	130	2	Q6ZSH0	Q6ZSH0	homo sapiens
12	59	27.3	130	2	BAC86983	Q9HT38	pseudomonas
13	59	27.3	134	2	Q9HT38	Q9HT38	pseudomonas
14	58	26.9	168	2	Q9SGN6	Q9SGN6	rhizobium l
15	57.5	26.6	82	2	Q853N5	Q853N5	mycobacteri
16	57.5	26.6	120	2	Q8T539	Q8T539	plasmodium
17	57	26.4	409	2	Q8C0D6	Q8C0D6	mus musculus
18	57	26.4	409	2	Q8C0F9	Q8C0F9	mus musculus
19	57	26.4	409	2	Q8C0L5	Q8C0L5	mus musculus
20	57	26.4	462	1	DCEA_LTSMO	Q9F5P3	listeria mo
21	57	26.4	464	1	DCBB_LTISIN	Q928I9	listeria mo
22	57	26.4	464	1	DCBB_LTSMO	Q928I9	listeria mo
23	57	26.4	464	2	Q8GF15	Q8GF15	edwardsiell
24	57	26.4	464	2	Q71X65	Q71X65	listeria mo
25	57	26.4	464	2	AA005100	AA005100	listeria mo
26	57	26.4	479	2	Q7V891	Q7V891	prochlorococ
27	57	26.4	560	2	Q7T0S1	Q7T0S1	xenopus lae
28	57	26.4	982	2	Q6WNK1	Q6WNK1	bdellovibri
29	57	26.4	982	2	CAE79150	CAE79150	bdellovibri
30	56.5	26.2	116	2	Q6ZF07	Q6ZF07	oryza sativ
31	56.5	26.2	116	2	BAC83424	BAC83424	oryza sativ

ALIGNMENTS

RESULT 1	ID	DCEA_ECOLI	STANDARD;	PRT;	466 AA.
AC	AT	P80063;			
AD	DT	01-WAR-1992 (Rel. 21, Created)			
AE	DT	01-DEC-1992 (Rel. 24, Last sequence update)			
AF	DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	DT	Glutamate decarboxylase alpha (EC 4.1.1.15) (GAD-alpha).			
GN	OS	Name=gadA; Synonyms=gads; OrderedLocusNames=B3517, C4328;			
OS	OS	Escherichia coli, and			
OS	OS	Escherichia coli O6.			
OC	OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	OC	Enterobacteriaceae; Escherichia.			
OX	OX	NCBI_TaxID=562, 217992;			
RT	[1]	SEQUENCE FROM N.A.			
RC	RC	STRAIN=K12;			
RC	RC	STRAIN=K12 / MG1655;			
RC	RC	MEDLINE=92394884; PubMed=1522060;			
RC	RC	Smith D.K., Kassam T., Singh B., Elliott J.F.;			
RT	RT	"Escherichia coli has two homologous glutamate decarboxylase genes			
RT	RT	that map to distinct loci."			
RL	RL	J. Bacteriol. 174:5820-5826(1992).			
RP	[2]	SEQUENCE, AND SEQUENCE OF 148-466 FROM N.A.			
RC	RC	STRAIN=ATCC 11246;			
RC	RC	MEDLINE=92155241; PubMed=1740158;			
RC	RC	Maras B., Sweeney G., Barra D., Bossa F., John R.A.;			
RT	RT	"The amino acid sequence of glutamate decarboxylase from Escherichia			
RT	RT	coli. Evolutionary relationship between mammalian and bacterial			
RT	RT	enzymes."			
RL	RL	Eur. J. Biochem. 204:93-98(1992).			
RP	[3]	SEQUENCE FROM N.A.			
RC	RC	STRAIN=K12 / MG1655;			
RC	RC	MEDLINE=94316500; PubMed=8041620;			
RC	RC	Sofia H.J., Eurland V., Daniels D.L., Plunkett G. III, Blattner F.R.			
RT	RT	"Analysis of the Escherichia coli genome. V. DNA sequence of the			
RT	RT	region from 76.0 to 81.5 minutes."			
RL	RL	Nucleic Acids Res. 22:2576-2586(1994).			
RP	[4]	SEQUENCE FROM N.A.			
RC	RC	STRAIN=O6.H1 / CFT073 / ATCC 700928 / UPSC;			
RC	RC	MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;			
RC	RC	Welch R.A., Eurland V., Plunkett G. III, Redford P., Roesch P.,			
RA	RA	Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.			
RA	RA	Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,			
RA	RA	McBey H.L.T., Donnenberg M.S., Blattner F.R.;			
RT	RT	"Extensive mosaic structure revealed by the complete genome sequence			
RT	RT	of uropathogenic Escherichia coli."			
RL	RL	Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).			
RP	[5]	SEQUENCE OF 1-22.			
RC	RC	MEDLINE=94033862; PubMed=7764225;			
RA	RA	Yoshida T., Yamashino T., Ueguchi C., Mizuno T.;			

RT "Expression of the Escherichia coli dimorphic glutamic acid
 RT decarboxylases is regulated by the nucleoid protein H-NS.";
 RL Biosci. Biotechnol. Biochem. 57:1568-1569(1993).
 RN [6]
 RP SEQUENCE OF 382-392.
 RC STRAIN=K12 / EMG2;
 RX MEDLINE=97443975; PubMed=9298646;
 RA Link A.J., Robison K., Church G.M.;
 RT "Comparing the predicted and observed properties of proteins encoded
 RT in the genome of Escherichia coli K-12.";
 RL Electrophoresis 18:1259-1313(1997).
 CC !- FUNCTION: Catalyzes the production of GABA.
 CC !- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
 CC !- COFACTOR: Pyridoxal phosphate.
 CC !- SUBUNIT: Homohexamer.
 CC !- SIMILARITY: Belongs to the group II decarboxylase family.
 CC !- DATABASE: NAME=Worthington enzyme manual;
 CC WWW="http://www.worthington-biochem.com/GLDP/".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M84024; AAA23833.1; -
 CC EMBL; X63123; CAA44834.1; -
 CC EMBL; U00039; AAB18493.1; -
 CC EMBL; A5000428; AAC76542.1; -
 CC EMBL; A5016768; AAN82764.1; ALT_INIT.
 CC F01; S47737; S24234.
 CC ECO2DBASE; D046.5; 6TH EDITION.
 CC ECO2DBASE; E046.5; 6TH EDITION.
 CC ECHOBASE; EB4302; -
 CC EcoGene; EGS0009; Gada.
 CC InterPro; IPR010107; Glu decarb GAD.
 CC InterPro; IPR002129; Pyridoxal dec.
 CC Pfam; PF00282; Pyridoxal dec; 1.
 CC TIGRPFAM; tigr01788; Glu-decarb-GAD; 1.
 CC PROSITE; P500392; DDC GAD_HDC_YDC; 1.
 CC Complete proteome; Decarboxylase; Direct protein sequencing; Lyase;
 KW Multigene family; Pyridoxal phosphate.
 KW BINDING 276 276
 FT BINDING 64 64
 FT CONFLICT 73 73 C -> S (in Ref. 2).
 FT CONFLICT 153 153 H -> R (in Ref. 2).
 FT CONFLICT 165 165 D -> N (in Ref. 2; AA sequence).
 FT CONFLICT 165 165 C -> S (in Ref. 2).
 FT CONFLICT 208 208 T -> N (in Ref. 2).
 FT CONFLICT 295 295 L -> V (in Ref. 2).
 FT CONFLICT 355 355 D -> N (in Ref. 2; AA sequence).
 SQ SEQUENCE 466 AA; 52685 MW; 86P963B710553E22 CRC64;
 Query Match 29.6%; Score 64; DB 1; Length 466;
 Best Local Similarity 50.0%; Pred. No. 4.7;
 Matches 14; Conservative 4; Mismatches 6; Indels 4; Gaps 1;
 QY 7 F0LRDLAD---AWDLSSRSRGRQWQP 30
 Db 380 FKLQGEDPGYTLDSLRLRGRQWQP 407
 RESULT 2
 ID DCEB_ECOLI STANDARD; PRT; 466 AA.
 AC P28302; P76873;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Glutamate decarboxylase beta (EC 4.1.1.15) (GAD-beta).
 GN Name=gadB; OrderedLocNames=b1493, 22215, EGS2098, SFI734, S1867;
 OS Escherichia coli,

OS Escherichia coli O157:H7, and
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562, 83334, 623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=K12;
 RX MEDLINE=92394884; PubMed=1522060;
 RA Smith D.K., Kassam T., Singh B., Elliott J.F.;
 RT "Escherichia coli has two homologous glutamate decarboxylase genes
 RT that map to distinct loci.";
 RL J. Bacteriol. 174:5820-5826(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=K12;
 RX MEDLINE=97251357; PubMed=9097039;
 RA Asai H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
 RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
 RA Sampei G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 28.0-40.1 min region on the linkage map.";
 RL DNA Res. 3:363-377(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
 RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai J., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Goddard E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
 RL Nature 409:529-533(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952 / EHEC;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kunara S., Shiba T., Hattori M., Shingawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 RN [6]
 RP SEQUENCE OF 1-318 FROM N.A.
 RC SPECIES=E.coli; STRAIN=K12;
 RA Turlin E., Gasser F., Biville F.;
 RT "Sequence and functional analysis of an Escherichia coli DNA fragment
 RT able to complement pqgE and pqgF from Methylobacterium organophilum.";
 RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 1-15.
 RC SPECIES=E.coli; STRAIN=K12;
 RX MEDLINE=93204884; PubMed=8455549;
 RA Yoshida T., Ueguchi C., Yamada H., Mizuno T.;
 RT "Function of the Escherichia coli nucleoid protein, H-NS: molecular
 RT analysis of a subset of proteins whose expression is enhanced in a hns

```

RT deletion mutant."
RL Mol. Gen. Genet. 237:113-122(1993).
RN [8]
RP SEQUENCE FROM N.A.
RX SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RC MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [9]
RP SEQUENCE FROM N.A.
RX SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
RC MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
CC -|- FUNCTION: Catalyzes the production of GABA.
CC -|- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
CC -|- COFACTOR: Pyridoxal phosphate.
CC -|- SUBUNIT: Homohexamer.
CC -|- SIMILARITY: Belongs to the group II decarboxylase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL; M84025; AAC23834.1; -.
DR EMBL; AE000246; AAC74566.1; -.
DR EMBL; D90791; BAA15163.1; -.
DR EMBL; D90790; BAA15157.1; -.
DR EMBL; AE005356; AAG56275.1; -.
DR EMBL; AP002557; AB35521.1; -.
DR EMBL; X71917; CAA50736.1; ALT SEQ.
DR EMBL; AE015194; AAN43309.1; ALT INIT.
DR EMBL; AE016984; AAP17196.1; -.
DR PIR; B43332; B43332.
DR PIR; B90891; B90891.
DR PIR; G85726; G85726.
DR PDB; 1PMM; X-ray; A/B/C/D/E/F=1-466.
DR EcoBASE; EB1453; -.
DR EcoGene; EG11490; gadB.
DR InterPro; IPR010107; Glu decarb GAD.
DR InterPro; IPR002129; Pyridoxal dec.
DR Pfam; PF00282; Pyridoxal dec; 1.
DR TIGRFAMs; TIGR01788; Glu-decarb-GAD; 1.
DR PROSITE; PS00392; DDC_GAD_HDC_YDC; 1.
DR 3D-structure; Complete proteome; Decarboxylase;
KW Direct protein sequencing; Lyase; Multi-gene family;
KW Pyridoxal phosphate.
FT BINDING 276 276 Pyridoxal phosphate.
SQ SEQUENCE 466 AA; 52668 MW; 8B653330A3C5B4ED CRC64;

Query Match 29.6%; Score 64; DB 1; Length 466;
Best Local Similarity 50.0%; Pred. No. 4.7;
Matches 14; Conservative 4; Mismatches 6; Indels 4; Gaps 1;

QY 7 FQLRDLDAD----AWDLSSRSRQRGWQMP 30
Db 380 FKLKDGEDPGTYLDLSERLRLRGWQVP 407

RESULT 4
Q83PR1 PRELIMINARY; PRT; 487 AA.
ID Q83PR1
AC Q83PR1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glutamate decarboxylase isozyme.
GN Name=gadA; OrderedLocusNames=SF3594;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."

```

```

RESULT 3
Q7UAY1 PRELIMINARY; PRT; 466 AA.
ID Q7UAY1
AC Q7UAY1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glutamate decarboxylase isozyme.
GN Name=gadA; OrderedLocusNames=S4173;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
CC -|- COFACTOR: Pyridoxal phosphate (By similarity).
CC -|- SIMILARITY: Belongs to the group II decarboxylase family.
DR EMBL; AE016991; AAP19142.1; -.
DR GO; GO:0016831; F:carboxy-lyase activity; IEA.
DR GO; GO:0008520; P:amino acid metabolism; IEA.
DR InterPro; IPR010107; Glu decarb GAD.
DR InterPro; IPR002129; Pyridoxal dec.
DR InterPro; IPR001638; SBP bac 3.
DR Pfam; PF00282; Pyridoxal dec; 1.
DR TIGRFAMs; TIGR01788; Glu-decarb-GAD; 1.
DR PROSITE; PS00392; DDC_GAD_HDC_YDC; UNKNOWN 1.
DR PROSITE; PS01039; SBP_BACTERIAL_3; UNKNOWN 1.
KW Lyase; Pyridoxal phosphate.
SQ SEQUENCE 466 AA; 52700 MW; 44B04876B2DF1680 CRC64;

Query Match 29.6%; Score 64; DB 2; Length 466;
Best Local Similarity 50.0%; Pred. No. 4.7;
Matches 14; Conservative 4; Mismatches 6; Indels 4; Gaps 1;

QY 7 FQLRDLDAD----AWDLSSRSRQRGWQMP 30
Db 380 FKLKDGEDPGTYLDLSERLRLRGWQVP 407

RESULT 4
Q83PR1 PRELIMINARY; PRT; 487 AA.
ID Q83PR1
AC Q83PR1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glutamate decarboxylase isozyme.
GN Name=gadA; OrderedLocusNames=SF3594;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."

```

```

RL Nucleic Acids Res. 30:4432-4441(2002).
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
CC -1- SIMILARITY: Belongs to the group II decarboxylase family.
DR EMBL; AE015367; AAN5045.1; -.
DR GO; GO:0016831; F:carboxy-lyase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR InterPro; IPR010107; Glu decarb GAD.
DR InterPro; IPR002129; Pyridoxal dec.
DR InterPro; IPR001638; SBP bac 3.
DR Pfam; PF00282; Pyridoxal dec 1.
DR TIGRFAMs; TIGR01788; Glu-decarb-GAD; 1.
DR PROSITE; PS00392; DDC_GAD_HDC_YDC; UNKNOWN 1.
DR PROSITE; PS01039; SBP_BACTERIAL_3; UNKNOWN 1.
KW Complete proteome; Lyase; Pyridoxal phosphate.
SQ SEQUENCE 487 AA; 55217 MW; 1B9D98041A7C38C2 CRC64;

Query Match 29.6%; Score 64; DB 2; Length 487;
Best Local Similarity 50.0%; Pred. No. 5;
Matches 14; Conservative 4; Mismatches 6; Indels 4; Gaps 1;

QY 7 FQLRDLDLAD----AWDLSSRSRQGWMP 30
DB 401 FKLKGDGDPGYTLSDLSERLRLRGWQP 428

RESULT 5
Q8FPHG5 PRELIMINARY; PRT; 489 AA.
AC Q8FPHG5;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Glutamate decarboxylase beta (ec 4.1.1.15).
GN Name=gadB; OrderedLocustNames=ec1922;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]

SEQUENCE FROM N.A.
STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G., III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA McInley H.L., Domenech M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
CC -1- SIMILARITY: Belongs to the group II decarboxylase family.
DR EMBL; AE016760; AAN80380.1; -.
DR GO; GO:0004351; F:glutamate decarboxylase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR InterPro; IPR010107; Glu decarb GAD.
DR InterPro; IPR002129; Pyridoxal dec.
DR Pfam; PF00282; Pyridoxal dec 1.
DR TIGRFAMs; TIGR01788; Glu-decarb-GAD; 1.
DR PROSITE; PS00392; DDC_GAD_HDC_YDC; UNKNOWN 1.
DR PROSITE; PS01039; SBP_BACTERIAL_3; UNKNOWN 1.
KW Complete proteome; Lyase; Pyridoxal phosphate.
SQ SEQUENCE 489 AA; 55440 MW; D885B97965EBCAF7 CRC64;

Query Match 29.6%; Score 64; DB 2; Length 489;
Best Local Similarity 50.0%; Pred. No. 5;
Matches 14; Conservative 4; Mismatches 6; Indels 4; Gaps 1;

QY 7 FQLRDLDLAD----AWDLSSRSRQGWMP 30
DB 403 FKLKGDGDPGYTLSDLSERLRLRGWQP 430

RESULT 6
O00833 PRELIMINARY; PRT; 234 AA.
ID O00833;
AC O00833;
DT 01-JUL-1997 (T-EMBLrel. 04, Created)
DT 01-JUL-1997 (T-EMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Trypanosoma cruzi.
OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5693;
RN [1]

SEQUENCE FROM N.A.
STRAIN=Cl Brener;
RX MEDLINE=20160916; PubMed=10688909;
RA Vazquez M., Ben-Dov C., Lorenzi H., Moore T., Schijman A.G.,
RA Levin M.J.;
RT "The short interspersed repetitive element of Trypanosoma cruzi, SIRE,
RT is part of VIPER, an unusual retroelement related to long terminal
RT repeat retrotransposons.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:2128-2133(2000).
RN [2]

SEQUENCE FROM N.A.
STRAIN=Cl Brener;
RA Vazquez M.P.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y12774; CAA73302.1; -.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 234 AA; 25905 MW; B900F6BB42652916 CRC64;

Query Match 28.0%; Score 60.5; DB 2; Length 234;
Best Local Similarity 44.8%; Pred. No. 6.7;
Matches 13; Conservative 3; Mismatches 12; Indels 1; Gaps 1;

QY 3 RGNFQRLDLADWL-SSRSRQGWMP 30
DB 201 RGRVFTQLDLAKGWLRRGAGSGWRTP 229

RESULT 7
Q8CBH3 PRELIMINARY; PRT; 269 AA.
ID Q8CBH3;
AC Q8CBH3;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length
DE enriched library, clone:9630027A13 product:RABSA, member RAS oncogene
DE family, full insert sequence. (Fragment).
GN Name=Rab5a;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]

SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]

SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TISSUE=Cerebellum;

```

RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630 (2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=20530913; PubMed=11079861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Saeki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Iehikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771 (2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nihi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Satoh K., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akai H., Tanaka Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK036011; BAC29275.1; -;
 DR MGD; MGI:105926; Rab5a.
 DR GO; GO:0005763; C:early endosome; IDA.
 DR GO; GO:0030139; C:cytosolic vesicle; IDA.
 DR GO; GO:0045121; C:lipid raft; IDA.
 DR GO; GO:0001726; C:ruffles; IDA.
 DR GO; GO:0003924; F:GTPase activity; IDA.
 DR GO; GO:0019001; F:guanyl nucleotide binding; IDA.
 DR GO; GO:0006897; P:apoptosis; IDA.
 FT NON TER 1
 SQ SEQUENCE 269 AA; 28676 MW; 1FF74B51B36340CD CRC64;
 Query Match 28.0%; Score 60.5; DB 2; Length 269;
 Best Local Similarity 50.0%; Pred. No. 7.8;
 Matches 12; Conservative 4; Mismatches 7; Indels 1; Gaps 1;
 QY 15 AWDLSRSRQGWQ-MNCRSRRG 37
 Db 144 AWDGNPLRRGHWPLPFRWRG 167
 RESULT 8
 Q6C917 PRELIMINARY; PRT; 220 AA.
 ID Q6C917
 AC Q6C917;
 DT 01-OCT-2004 (TrEMBLrel. 28, Created)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DE 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Similar to sp|P21374|Saccharomyces cerevisiae YJR050w UTR3 pre-mRNA
 DE splicing factor.

GN ORFNames=YALI0D109019;
 OS Yarrowia lipolytica (Candida lipolytica).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Dipodascaceae; Yarrowia.
 OX NCBI_TaxID=4952;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIB99;
 RG GENOEVUES;
 RA Dujon B., Sherman D., Fischer G., Durkens P., Casaregola S.,
 RA Lafontaine I., de Montigny J., Marck C., Neveglisse C., Talla E.,
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
 RA Barney S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
 RA Boistrame A., Boyer J., Cattolico L., Confanioulet F., de Daruvar A.,
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
 RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
 RA Niclaud J.M., Nikolski M., Ostas S., Olier-Kalogeropoulos C.,
 RA Pellens S., Potier S., Richard G.F., Straub M.L., Suleau A.,
 RA Swemene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
 RA Zeniou-Meyer M., Zivancovic I., Bolotin-Fukuhara M., Thierry A.,
 RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
 RA Wincker P., Souciet J.L.;
 RT "Genome evolution in yeasts.";
 RL Nature 430:35-44 (2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIB99;
 RA Genoscope;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CR382130; CAG80863.1; -; B172423C76C30B7B CRC64;
 SQ SEQUENCE 220 AA; 25196 MW; B172423C76C30B7B CRC64;
 Query Match 27.8%; Score 60; DB 2; Length 220;
 Best Local Similarity 33.9%; Pred. No. 7.3;
 Matches 19; Conservative 6; Mismatches 9; Indels 22; Gaps 3;
 QY 5 NEFQDLAD-----AWDLSSRS-----RQWQNPNCRRGPG 39
 Db 70 SDFQRLDNLNKLMSERHWDQLIRSLGGPNFYASFGCKRGYQYVG-RARELPG 124
 RESULT 9
 DCEA_ECO57
 ID DCEA_ECO57 STANDARD; PRT; 466 AA.
 AC P58228;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Glutamate decarboxylase alpha (EC 4.1.1.15) (GAD-alpha).
 GN Name-gada; Synonyms=gads; OrderedLocusNames=z4930, EC44397;
 OS Escherichia coli O157:H7.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
 RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouisis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
 RL Nature 409:529-533 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952 / EHEC;
 RX MEDLINE=21156211; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,

RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.,
 RA "Complete genome sequence of enterohemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12."
 RT DNA Res. 8:111-22(2001).
 CC -!- FUNCTION: Catalyzes the production of GABA (By similarity).
 CC -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
 CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
 CC -!- SUBUNIT: Homohexamer (By similarity).
 CC -!- SIMILARITY: Belongs to the group II decarboxylase family.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB005577; AAG58658.1; -;
 CC EMBL; AF002565; BAB37820.1; -;
 CC PIR; E91178; E91178.
 CC PIR; F86024; F86024.
 CC InterPro; IPR010107; Glu decarb GAD.
 CC InterPro; IPR002129; Pyridoxal dec.
 CC Pfam; PF00282; Pyridoxal dec. 1.
 CC TIGRFAMs; TIGR01788; Glu-decarb-GAD; 1.
 CC PROSITE; P500392; DDC GAD HDC YDC; 1.
 CC Complete proteome; Decarboxylase; Lyase; Multigene family;
 KW Pyridoxal phosphate.
 FT BINDING 276 Pyridoxal phosphate (By similarity).
 SQ SEQUENCE 466 AA; 52699 MW; C7F9623DBE24E489 CRC64;

Query Match 27.8%; Score 60; DB 1; Length 466;
 Best Local Similarity 46.4%; Pred. No. 16;
 Matches 13; Conservative 5; Mismatches 6; Indels 4; Gaps 1;

QY 7 FQRLRLAD----AWLSSRSRQRGWMP 30
 DB 380 FKLKEGEDPGYTLDSLRLRQGWQVP 407

RESULT 10
 ID Q81VF8 PRELIMINARY; PRT; 730 AA.
 AC Q81VF8;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE KIAA2013 protein (Fragment).
 GN Name-KIAA2013.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Nagase T., Kikuno R., Ohara O.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB095933; BAC23109.1; -;
 FT NON_TER 1
 SQ SEQUENCE 730 AA; 79351 MW; 0D048672E371AEC8 CRC64;

Query Match 27.5%; Score 59.5; DB 2; Length 730;
 Best Local Similarity 46.2%; Pred. No. 31;
 Matches 12; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

QY 14 DAWDLSSRSRQRGWQMPNCRSRGPG 39
 DB 3 DGRWPPRRRRKQGWRLRLRLRRGPG 27

RESULT 11
 ID Q6ZSH0 PRELIMINARY; PRT; 130 AA.
 AC Q6ZSH0;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein FLJ45539.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thalamus;
 RA Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Sugiyama T.,
 RA Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J.,
 RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
 RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
 RA Masuho Y., Nagai K., Isogai T.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK127447; BAC86983.1; -;
 SQ SEQUENCE 130 AA; 14426 MW; 4FE3D779B71989AF CRC64;

Query Match 27.3%; Score 59; DB 2; Length 130;
 Best Local Similarity 32.1%; Pred. No. 5.7; Indels 0; Gaps 0;
 Matches 9; Conservative 5; Mismatches 14;

QY 5 NEFQLRLADAWDLSSRSRQRGWQMPNC 32
 DB 19 NSLEAPYLVSGWGWGSRERKQWELITC 46

RESULT 12
 ID BAC86983 PRELIMINARY; PRT; 130 AA.
 AC BAC86983;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE CDNA FLJ45539 fis, clone BRTHA203122.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thalamus;
 RA Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Sugiyama T.,
 RA Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J.,
 RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
 RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
 RA Masuho Y., Nagai K., Isogai T.;
 RL "NEDO human cDNA sequencing project."
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK127447; BAC86983.1; -;
 SQ SEQUENCE 130 AA; 14426 MW; 4FE3D779B71989AF CRC64;

Query Match 27.3%; Score 59; DB 2; Length 130;
 Best Local Similarity 32.1%; Pred. No. 5.7; Indels 0; Gaps 0;
 Matches 9; Conservative 5; Mismatches 14;

QY 5 NEFQLRLADAWDLSSRSRQRGWQMPNC 32
 DB 19 NSLEAPYLVSGWGWGSRERKQWELITC 46

RESULT 13
 Q9HT38

```

ID Q9HT38 PRELIMINARY; PRT; 134 AA.
AC Q9HT38;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PA5536;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAOI;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,
RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Golltry L.N., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.B.W., Lory S., Olson K.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
RT opportunistic pathogen."
RL Nature 406:959-964 (2000).
DR EMBL; AE004966; AAG08921.1; --
DR PIR; B92953; E82953.
DR InterPro; IPR000962; Znf_DsKa/TrAr.
DR Pfam; PF01258; zf-dsKa_c1ar; 1.
DR PRINTS; PR00618; DKSZNFINGER.
DR Complete Proteome; Hypothetical protein.
KW SEQUENCE 134 AA; 16043 MW; 88E8A7F687B692A0 CRC64;

Query Match 27.3%; Score 59; DB 2; Length 134;
Best Local Similarity 53.6%; Pred. No. 5.9;
Matches 15; Conservative 4; Mismatches 5; Indels 4; Gaps 2;

QY 6 EF-QLRDL---ADAWDLSSRSRQRGQM 29
|||:||||:|:|||||:|:|:
Db 41 EFGELRDLRPSDEADLASREQRWQL 68

RESULT 14
Q98GN6 PRELIMINARY; PRT; 168 AA.
ID Q98GN6
AC Q98GN6;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ml13243 protein.
GN OrderedLocusNames=ml13243;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082936; PubMed=11214974;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti (supplement).";
RL DNA Res. 7:381-406 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,

```

```

RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338 (2000).
DR EMBL; AP003001; BAB50180.1; --
DR HSP; P32939; IKY3.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; Ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
KW Complete proteome; GTP-binding.
SQ SEQUENCE 168 AA; 18516 MW; 47F9B29D3EF0850 CRC64;

Query Match 26.9%; Score 58; DB 2; Length 168;
Best Local Similarity 44.4%; Pred. No. 10;
Matches 12; Conservative 3; Mismatches 8; Indels 4; Gaps 1;

QY 7 FOURDLADAWDLs---SRSRQRGQM 29
|:|||||:|:|:|:|:|:
Db 118 FNKSLADRWELSDGIDNLRQLGWQI 144

RESULT 15
Q853N5 PRELIMINARY; PRT; 82 AA.
ID Q853N5
AC Q853N5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp42.
OS Mycobacteriophage Bx1.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
CX NCBI_TaxID=205877;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22592660; PubMed=12705866;
RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,
RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,
RA Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,
RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
RA Hatfull G.F.;
RT "Origins of highly mosaic mycobacteriophage genomes.";
RL Cell 113:171-182 (2003).
DR EMBL; AY129337; AAN16702.1; --
SQ SEQUENCE 82 AA; 9442 MW; 9D58148FEA7486EC CRC64;

Query Match 26.6%; Score 57.5; DB 2; Length 82;
Best Local Similarity 48.6%; Pred. No. 5.6;
Matches 17; Conservative 1; Mismatches 12; Indels 5; Gaps 4;

QY 6 EFQLRDL-ADAWDLSSRSRQRGQM--PNCRRRG 37
|:|||||:|:|:|:|:|:
Db 49 EAAEDLRDGN-LVLRER-RGWSLICPNCRREG 81

Search completed: November 10, 2004, 14:50:34
Job time : 53.9509 secs

```

This Page Blank (uspio)